

# Social transformations and labour market entry

Citation for published version (APA):

Rossello, G. (2021). *Social transformations and labour market entry: an investigation into university systems in emerging economies*. [Doctoral Thesis, Maastricht University]. ProefschriftMaken. <https://doi.org/10.26481/dis.20210202gr>

**Document status and date:**

Published: 01/01/2021

**DOI:**

[10.26481/dis.20210202gr](https://doi.org/10.26481/dis.20210202gr)

**Document Version:**

Publisher's PDF, also known as Version of record

**Please check the document version of this publication:**

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

**General rights**

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

[www.umlib.nl/taverne-license](http://www.umlib.nl/taverne-license)

**Take down policy**

If you believe that this document breaches copyright please contact us at:

[repository@maastrichtuniversity.nl](mailto:repository@maastrichtuniversity.nl)

providing details and we will investigate your claim.

# **Social Transformations and Labour Market Entry**

An Investigation into University Systems in  
Emerging Economies

by Giulia Rossello

ISBN 978-94-6423-117-5

Copyright © 2021 by Giulia Rossello

Cover design: Giulia Rossello and ProefschriftMaken

Publisher: ProefschriftMaken, Maastricht, The Netherlands

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without the prior permission in writing, from the author.

# **Social Transformations and Labour Market Entry**

An Investigation into University Systems in  
Emerging Economies

Dissertation

to obtain the degree of Doctor at Maastricht University, on the authority of the Rector  
Magnificus, Prof. Dr. Rianne M. Letschert, in accordance with the decision of the Board of  
Deans, to be defended in public on Tuesday 2nd February 2021, at 16.00 hours

Giulia Rossello



**Supervisor:** Prof. Robin Cowan

**Co-Supervisor:** Prof. Jacques Mairesse

**Assessment Committee:**

Prof. Dr. Eleonora Nillesen (Chair)

Dr. Heidi Prozesky (CREST, Centre of Excellence in Scientometrics and Science, Technology and Innovation Policy, Stellenbosch University, South Africa)

Prof. Dr. Paula Stephan (Andrew Young School of Policy Studies, Georgia State University, United States)

Dr. Fabiana Visentin

*Dedicato a tutte le persone che mi hanno ispirato  
ed aiutato in questo percorso.*



# Summary

A demographic balance in desirable occupations is socially important. Compared to white males, women and people of colour tend to have lower labour market outcomes and to be under-represented in top job positions. Despite the increasing participation of people of colour and females in higher education, few become faculty members in universities. The inclusion of previously excluded groups in academia is particularly important for developing countries that experienced large social transformations and have less mature knowledge systems. Even if gender and racial unbalances have been extensively investigated, many studies overlook the fundamental characteristics of the process of entry into academia. In particular, this dissertation examines how individual and systemic mechanisms relate with supervision, collaboration, and mutual evaluation when they first manifest in an academic career.

Mentorships and supervision are known to affect the students' future careers development and involve one of the first academic tie for a student. Therefore, the second chapter examines the formation of student-advisor relationship and whether race and gender homophily plays a role in the tertiary education. The chapter uses student enrolment data of bachelor, master, and Ph.D. students and their advisors and measures the relative magnitudes of induced and choice homophily race- (gender-) based homophily. Further, the chapter examines where choice homophily originates, estimating homophilous preferences in the interacting populations of black and white (male and female) students and advisors.

The results of chapter 2 highlight that the process of racial and gender inclusion is still in a transition phase. Indeed, both race-based and gender-based homophily play a crucial role in tie formation. In particular, the university system itself induces half of the observed homophily levels while half arises from individual choices. The latter mostly originates in the population of white/male students and black/female advisors. Further, the chapter highlights the potential tension between social transformation and effective inclusion. Even where homophilous preferences are unchanged

at the individual level, a demographic change in the composition of the population can mechanically increase the aggregate homophily in the system and, in turn, slow down the inclusion.

The third chapter looks at the training phase, analysing the first co-authorship link of students — the one with their doctoral advisors. Doctoral training is critical today because job market committees often use candidate publications during the Ph.D. as a screening criterion. Starting from these premises, chapter three studies whether homophily and advisor characteristics influence students' productivity during their doctorates. In particular, the chapter asks whether the (racial and/or gender) composition of student-advisor matter for publication productivity of Ph.D. students and whether the relationship varies for different student productivity-profiles.

The chapter finds that females working with a female advisor are as productive as male students working with a male advisor. While in contrast, when female students work with a male, they are less productive compared to male-male student-advisor couples. Interestingly, the gender difference is u-shaped over student productivity and cross-racial ties display almost no gender difference in productivity. Indeed, looking the sub-samples of white-white and black-black, couples gender differences in average productivity across the groups (and pairs) are driven by differences in the upper tail of the productivity distribution.

Chapters four and five go further to investigate the transition from the Ph.D. to the first academic job using data from South Africa and Mexico. In particular, the chapters study the link between university prestige stratification and research performance in the short, medium, and long-run. The chapters highlight predictable hiring patterns from one university to another that follow an underlying prestige hierarchy. Moreover, in both systems, internal hiring of Ph.D. graduates is positively related to academic performance. This underlines a positive role of inertia that might prevent talents and knowledge exchanges among universities.

This dissertation examines process of entry into the workplace looking at individual and systemic mechanisms. Globally, this work underlines how homophily, Matthew effects, and prestige stratification are strong mechanisms in the workplace. These mechanisms can inhibit social transformations at the aggregate level and at the same time be beneficial for individual performance. This ambiguity might explain why transformations are often unevenly distributed in the job hierarchy, fields, and universities. This book highlights the need for continuous policy intervention and multiple targets to achieve sustainable social transformations in the workplace.

*“I have dedicated my life to this struggle of the African people. I have fought against white domination, and I have fought against black domination. I have cherished the ideal of a democratic and free society in which all persons live together in harmony with equal opportunities. It is an ideal which I hope to live for, and to see realized. But if needs be, it is an ideal for which I am prepared to die.”*

— Nelson Mandela

Speech at his trial in Pretoria, 20 April 1964, which he quoted on his release in Cape Town, 11 February 1990

# Acknowledgements

It is my pleasure to acknowledge the people who have been crucial for the completion of my research during the Doctorate. First and foremost, my thanks go to my supervisor, Prof. Robin Cowan. His role was crucial; he was able to guide me and to stimulate my passion for research and creativity. He was patient with my, sometimes, rude English. He was always able to understand my thoughts also when my formalisation was tacit. Working with him was an incentive to improve and develop myself. He is undoubtedly a role model for me. I am also indebted to Prof. Jacques Mairesse, my second supervisor, who guided me with his experience and knowledge. He always provided new challenges, comments, and support. I would like to thank the members of the assessment committee Prof. Eleonora Nillesen, Dr. Heidi Prozesky, Prof. Paula Stephan, and Dr. Fabiana Visentin whose valuable comments have enriched this thesis. I want to thank also Prof. Moritz Mueller, the faculty and staff of BETA at Strasbourg University (France) and of the CREST institute at Stellenbosh University (South Africa). I want to thank Prof. Giorgio Fagiolo and Prof. Alessandro Nuvolari, who convinced me to start this Ph.D. in the first place and professor Arianna Martinelli for the opportunity she gave me to continue research. Working at the faculty of UNU-MERIT was a pleasure and the possibility to work and know colleagues from all over the world thought me a lot about the world. So I thank all the administrative and support staff, they always work to make everything organised and planned thinking of nice parties and celebration occasions too. Above all a very spe-

cial thank goes to Eveline, she was “the light in the darkness of winter” and her office was “the home abroad”. My years in the Netherlands would not have been the same without all the friends who have accompanied me during the journey. I was lucky to have a great batch Beatrix, Rose, Patima, Tamara, Halefom, Danilo, Emmanuel, Victor, Joseph, Kaleab, and Bart thanks for all the laughs. A very special thank goes to Michelle G., Mario, Francesca, Hugo, Ibrahima, Elisa, Jennifer, Michelle M., Sheng, Racky, Flavia, Juanca, Allison, Marijo, Oriol, Mary, Hiwot, Mueid, Gintare, Solomon, Maisha, Sabrina, Robert, Maty, Davina, Gideon, Charlotte, Tatenda, Simone, Daniel, Lorena, Florencia, and Caio. Then I want to thank the friends I met in Pisa. I have neglected them over the Maastricht years, but that they were with me over many adventures Sophie, Giulia, Carla, Francesca, Alessandra, Annarita, Simone, Fabio, Silvia, Tazio, Donald, Albara, Marco, Beatrice, Silvia, Carlo, Adriana, Emanuele, May, Francesco, Alessandro, Marica, Jacopo, and Matteo. Finally, I want to thank my family. Without the encouragement and support of my mum and dad none of this would have been possible. Moreover, I want to thank my sister for complementing me with her art and her grumbles. Last but not least, the most important person in my life, the family I choose, my life partner Daniele. Thank you for all your support for always believing in me also when I doubt myself. I would have never learned to fly without you.

*Maastricht, Academic year 2020-2021*

G. R.

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Entering process . . . . .	6
1.1.1	Network (or Systemic) mechanisms of the entry process . . . .	6
1.2	Entry in a job in academia . . . . .	11
1.2.1	Homophily in tie formation . . . . .	13
1.2.2	Doctoral productivity . . . . .	15
1.2.3	First job . . . . .	17
1.3	Social Transformations and the Entry Process . . . . .	18
1.3.1	Mechanisms that foster or hinder social transformations . . . .	18
1.4	The Context . . . . .	20
1.4.1	South Africa . . . . .	20
1.4.2	Mexico . . . . .	22
1.5	Knowledge Gap . . . . .	23
1.6	Methods . . . . .	26
1.7	Thesis outline . . . . .	27
<b>2</b>	<b>Homophily in student-supervisor tie formation</b>	<b>29</b>
2.1	Introduction . . . . .	30
2.2	Homophily and social transformations . . . . .	32
2.2.1	Homophily in Higher Education and Research . . . . .	34
2.3	Data . . . . .	36
2.4	Preliminaries . . . . .	38
2.4.1	Null Model . . . . .	38
2.4.2	Homophily and Assortativity . . . . .	39
2.5	Student-Supervisor ties . . . . .	40
2.5.1	Future Academics . . . . .	41
2.6	Methodology . . . . .	44
2.7	Permutation Test . . . . .	45



2.7.1	Permutation Test results . . . . .	46
2.8	Choice and Induced homophily . . . . .	49
2.9	Where does choice homophily originate? . . . . .	52
2.9.1	A Model of homophilous preferences of the different groups . . . . .	53
2.9.2	Choice Homophily of different groups . . . . .	54
2.10	A simple model with changing populations and unchanging homophily . . . . .	56
2.11	Conclusion . . . . .	57
<b>3</b>	<b>Doctoral productivity and the role of gender and race in supervision</b>	<b>59</b>
3.1	Introduction . . . . .	60
3.2	Material . . . . .	65
3.2.1	Data . . . . .	65
3.2.2	Variable description . . . . .	67
3.3	Methods . . . . .	68
3.4	Results . . . . .	72
3.4.1	Results looking at race . . . . .	75
3.5	Conclusion . . . . .	75
<b>4</b>	<b>Emergent structures in faculty hiring networks, and the effects of early-career mobility on academic performance</b>	<b>79</b>
4.1	Introduction . . . . .	80
4.2	University prestige and young faculty hiring . . . . .	82
4.3	Data and variable construction . . . . .	87
4.3.1	Faculty hiring network . . . . .	88
4.3.2	Prestige ranking . . . . .	90
4.4	Effects of rank change on future career . . . . .	94
4.4.1	Matched pairs . . . . .	94
4.5	Results . . . . .	95
4.5.1	Possible cohort effects . . . . .	100
4.5.2	Size effects . . . . .	100
4.6	Discussion and Conclusion . . . . .	100
<b>5</b>	<b>University prestige in the Ph.D. job market and the big-fish-little-pond effect</b>	<b>105</b>
5.1	Introduction . . . . .	106
5.2	Prestige, mobility and performance . . . . .	108
5.2.1	University prestige . . . . .	108
5.2.2	The link between mobility and performance . . . . .	109
5.3	Data . . . . .	110

5.4	Interactive Prestige Ranking . . . . .	111
5.5	Prestige Ranking Results . . . . .	113
5.6	Matched Pairs Analysis . . . . .	116
5.7	Matched Pairs Results . . . . .	117
5.7.1	Robustness check . . . . .	120
5.8	Discussion . . . . .	122
5.9	Conclusion . . . . .	125
<b>6</b>	<b>Conclusion</b>	<b>127</b>
	<b>Appendices</b>	<b>132</b>
A.1	Permutation Test results, Choice and Induced Homophily . . . . .	133
A.2	List of Institutions and Fields in the data . . . . .	134
A.3	Robustness check . . . . .	135
A.4	Gender Analysis . . . . .	137
A.4.1	Future Academics . . . . .	137
A.4.2	Permutation Test . . . . .	138
A.4.3	Choice and Induced Homophily . . . . .	140
A.4.4	Choice Homophily of different groups . . . . .	142
B.5	Variables . . . . .	145
B.6	Additional Statistics on the data . . . . .	145
B.7	Regression with gender and race interaction on the whole sample . . .	152
B.8	Main OLS regressions . . . . .	152
B.9	Poisson panel regressions Gender . . . . .	154
B.9.1	Students . . . . .	154
B.9.2	Advisors . . . . .	155
B.9.3	Student and Advisor couple . . . . .	156
B.10	Quantile Regressions Gender . . . . .	157
B.11	Race Results . . . . .	159
B.12	Poisson panel regressions race . . . . .	162
B.12.1	Students . . . . .	162
B.12.2	Advisors . . . . .	163
B.12.3	Student and Advisor couple . . . . .	164
B.13	Quantile Regressions Race . . . . .	165
B.14	Without controls . . . . .	167
C.15	University reform in 2004 . . . . .	171
C.16	Faculty hiring network . . . . .	173
C.17	Prestige Ranking aggregating SET and SSH . . . . .	174

C.18 Cohort effects . . . . .	175
C.19 SSH results . . . . .	179
C.20 KS test matched pairs . . . . .	182
C.21 Robustness check of KS test . . . . .	185
D.22 Faculty Hiring Matrix Names . . . . .	189
D.23 Prestige Ranking Algorithm . . . . .	190
D.24 NSR: Disciplines, Evaluation and Commissions . . . . .	191
D.25 Results Static Rank . . . . .	192
D.26 KS-Test Results . . . . .	195
<b>Valorisation Addendum</b>	<b>216</b>
<b>About the Author</b>	<b>222</b>

# List of Figures

2.1	Stack plot of Student-Supervisor relations 1973-2014. . . . .	41
2.2	Student-Supervisor network for Future Academics 1973-2014. . . . .	43
2.3	Permutation tests for student supervisor data 1973-2014. . . . .	48
2.4	Induced and Choice Homophily comparison looking racial ties in student supervisor relations 1973-2014. . . . .	50
2.5	Race Homophily of the different groups 1973-2014. . . . .	54
2.6	Model predictions as a function of number of black in the population .	57
3.1	Heat-map of doctoral average annual productivity for student and advisor gender (racial) combinations. . . . .	67
3.2	Quantile Regressions for student annual average doctoral productivity comparing student-advisor gender couples. . . . .	73
4.1	Hiring networks 1970-2004 SET. . . . .	89
4.2	Prestige Ranking for SET 1970-2004. . . . .	92
4.3	Up versus stay comparison. . . . .	96
4.4	Down versus stay comparison. . . . .	97
4.5	Up versus down comparison. . . . .	99
5.1	Up versus Stay comparison. . . . .	118
5.2	Down versus Stay comparison . . . . .	119
5.3	Down versus Stay comparison. . . . .	120
1	Permutation test for student supervisor data 1973-2014 for top institutions and Phd students. . . . .	133
2	Induced and Choice Homophily comparison looking racial ties in student supervisor relations 1973-2014. . . . .	134
3	Student-Supervisor network for Future Academics 1973-2014. . . . .	137
4	Permutation test for student supervisor data 1973-2014. . . . .	138

5	Permutation test for student supervisor data 1973-2014 for top institutions and PhD students. . . . .	139
6	Induced and Choice Homophily comparison looking gender ties in student supervisor relations 1973-2014. . . . .	140
7	Induced and Choice Homophily comparison looking gender ties in student supervisor relations 1973-2014. . . . .	141
8	Gender Homophily of the different groups 1973-2014. . . . .	142
9	Average three years publications of students and advisors . . . . .	145
10	Cumulative number of publications of advisor and couples . . . . .	147
11	Three years number of publications of students and couples . . . . .	150
12	Distribution of the dependent variable by student-advisor gender couple	151
13	Quantile Regression with clustered standard errors. . . . .	160
14	Prestige Ranking 1970-2004 without distinction of fields. . . . .	174
15	Up versus stay comparison, PhD degrees prior to 1992. . . . .	176
16	Down versus stay comparison, PhD degrees prior to 1992. . . . .	177
17	Up versus down comparison, PhD degrees prior to 1992. . . . .	178
18	Hiring network 1970-2004 SSH. . . . .	180
19	Prestige Ranking for SSH 1970-2004. . . . .	182
20	Stochastic Analysis of Up vs Stay. . . . .	192
21	Stochastic Analysis of Stay vs Down . . . . .	193
22	Stochastic Analysis of Stay vs Down . . . . .	194

# List of Tables

2.1	Student supervision composition. . . . .	40
2.2	Student supervision composition subsample of Future Academics. . .	42
3.1	Students and Advisors, by Race and Gender. . . . .	66
3.2	Pooled OLS Panel Regression with robust clustered standard error. . .	71
4.1	Adjacency matrix of the hiring network for the years 1970-2004 . . .	88
4.2	Summary Statistics SET hiring network for the years 1970-2004. . . .	90
4.3	PhDs hired from the top 5 prestigious universities in SET . . . . .	93
5.1	Ranking Mexican Universities 1992-2016 for STEM Sciences. . . . .	114
5.2	Ph.D. to First-Job Mexican Faculty Hiring Matrix 1992-2016 for STEM Sciences. . . . .	114
5.3	Prestige movements per university prestige 1992-2016 . . . . .	115
5.4	Ordered logistic regressions: Impact of prestige changes on academic performance. . . . .	121
1	Results of logistic regressions of yearly student-supervisor racial tie data. . . . .	136
2	Results of logistic regressions of yearly student-supervisor gender tie data . . . . .	143
3	Student 3 year average number of publications . . . . .	146
4	Advisor Logarithm of 1+ cumulative average productivity from first record to t-1 . . . . .	146
5	Distribution of the study sample by year of thesis defence . . . . .	147
6	Distribution of Ph.D. students and advisors by the discipline of the thesis. . . . .	148
7	Average productivity for different sub-sample of the data and student advisor couple. . . . .	148

8	Gender and Racial Assortativity Coefficient (Ass.) by universities and field. . . . .	149
9	Pooled OLS Panel Regression on the whole sample with interaction terms . . . . .	152
10	Pooled OLS Panel Regression with robust clustered standard error (Main Model). . . . .	153
11	Poisson Panel Regression with robust clustered standard error (students).154	
12	Poisson Panel Regression with robust clustered standard error (advisors).155	
13	Poisson Panel Regression with robust clustered standard error (couples).156	
14	Quantile regression with clustered standard errors on the complete sample. . . . .	157
15	Quantile regression with clustered standard errors on the sub-sample of White Student and White Advisor . . . . .	157
16	Quantile regression with clustered standard errors on the subsample of Black Student and Black Advisor. . . . .	158
17	Quantile regression with clustered standard errors on the subsample of Black Student and White Advisor. . . . .	158
18	Pooled OLS Panel Regression with robust clustered standard error (Main model). . . . .	159
19	Quantile regression with robust clustered standard errors . . . . .	161
20	Poisson Panel Regression with robust clustered standard error (students).162	
21	Poisson Panel Regression with robust clustered standard error (advisors).163	
22	Poisson Panel Regression with robust clustered standard error (couples).164	
23	Quantile regression with clustered standard errors on the whole sample. 165	
24	Quantile regression with clustered standard errors on the sub-sample of Male Student and Male Advisor . . . . .	165
25	Quantile regression with clustered standard errors on the sub-sample of Female Student and Female Advisor. . . . .	166
26	Quantile regression with clustered standard errors on the sub-sample of Female Student and Male Advisor. . . . .	166
28	Pooled Panel Regression without controls with robust clustered standard error (advisors) . . . . .	167
27	Pooled Panel Regression without controls with robust clustered standard error (students). . . . .	167
29	Pooled Panel Regression without controls with robust clustered standard error (couples) . . . . .	168
30	Pooled Panel Regression without controls with robust clustered standard error (students) . . . . .	168

31	Pooled Panel Regression without controls with robust clustered standard error (advisors) . . . . .	169
32	Pooled Panel Regression without controls with robust clustered standard error (couples) . . . . .	169
33	Prestige Ranking with pre-merger names . . . . .	172
34	Indegree and Outdegree SET hiring network for the years 1970-2004. . . . .	173
35	Adjacency matrix of the hiring network for the years 1970-2004 in SSH . . . . .	179
36	Summary Statistics SSH hiring network for the years 1970-2004. . . . .	181
37	PhDs hired from the top 5 prestigious universities in SSH . . . . .	181
38	Results of KS test relative to figure 4.3, 4.4, and 4.5. . . . .	183
39	Results of KS test relative to figures 15, 16, and 17. . . . .	184
40	Proportion of p-values lower than 0.05 of Bootstrap KS test repeated 100 times on repeated re-samples of size=100 for p% . . . . .	186
41	Proportion of p-values lower than 0.05 of Bootstrap KS test repeated 100 times on repeated re-samples of size=100 for p%. PhD degrees prior than 1992 . . . . .	187
42	KS-Test Results for Phd and First-job Matches. . . . .	195





# Chapter 1

## Introduction

Excluded groups are everywhere in the world. Compared to white males, women and minority groups tend to have lower labour market outcomes, have their jobs concentrated in specific occupations and are under-represented in top job positions. Examining 84 countries, the recent report ILO-UN Women (2019) finds that half of the female population is excluded or marginalised from the labour market, having a labour force participation rate 43 percentage points lower than males.<sup>1</sup> Besides significant regional differences, the situation is problematic also for rich countries where labour markets statistics vary for demographic groups. The recent social tension in the U.S. and the growing activism of Black Lives Matters testify an unsustainable situation. The 2018 report of the U.S. Economic Policy Institute finds that African American workers have an unemployment rate of 7.2 percent, which is the highest compared to other demographic groups.<sup>2</sup> Moreover, the U.S. Bureau of Labor Statistics (BLS) records a distressing trend, where, between the year 2000 and 2016, the participation of African Americans to the labour force declines from 65.8 percent to 61.6 percent. In addition, the McKinsey report “*Women in the workplace*” (2019) who covers 590 companies and more than 22 million employees, highlights that “*Women and people of colour remain under-represented at every level in corporate America, despite earning more college degrees than white men for thirty years and counting*”.

This waste of talents and human capital causes incalculable societal, cultural, knowledge, and economic losses in our societies. How many ideas have been wasted? How much advice has been unheeded? Which investment decisions would have been

---

<sup>1</sup>See <https://ilostat.ilo.org/having-kids-sets-back-womens-labour-force-participation-more-so-than-getting-married/>; last access August 2020

<sup>2</sup>See <https://www.epi.org/publication/state-race-unemployment-2018q1/> Janelle Jones's report *Economic Indicators*, commissioned by the Economic Policy Institute in 2018

taken from a more demographically balanced firm board? How would industrial policies have evolved if more blacks and females were in the decision rooms? Reversing the problem, one can try to quantify the contribution that inclusion has brought to GDP per capita and economic growth. Hsieh et al. (2013) estimate that the decreasing barriers for blacks and women observed in the U.S. between 1960 and 2010 explain 24% of growth in GDP per capita and 6% of economic growth. But besides any attempt to quantify gains and losses, history offers many quasi-natural experiments where inclusion was accelerated by exogenous conditions. During WWI and WWII, women massively entered the U.S. and the European labour markets (Hart, 2009; Bellou and Cardia, 2016; Gay, 2019; Boehnke and Gay, 2020). Their involvement contributed to the sustained period of growth and innovation post WWII. In this respect, looking at U.S. data Goldin and Olivetti (2013) find that WWII provoked an increase of the fraction of woman workers especially among the highly educated women that were previously excluded by social norms. Similar examples are also found in developing contexts. For example, in Rwanda, after the genocide of 1994 (that killed many of the men), women rebuilt the country. Since then, the country is increasing its per capita GDP at extraordinary rates, and many argue that this is mostly driven by the commitment of the Rwandan government to achieve gender equality (Abbott and Malunda, 2016).<sup>3</sup> However, the work of Abbott and Malunda (2016) also examines recent shadows. They highlight that the effort towards women's empowerment significantly decreased after 2003 and it is not uniformly distributed in the country across job-ranks and geographical areas.

Lights and shadows of those examples underline that achieving a balanced representation of the population is socially and economically important. In recent decades many countries have seen improvements in the participation of female and people of colour in many occupations. Indeed, women and blacks made large investments in their higher education (Saleh et al., 2017; Allen, 2018; Blalock and Sharpe, 2012; Gemici and Wiswall, 2014), and those investments in human capital changed the composition of the student body at universities in many countries. Today, women and people of colour are represented as students in art, history, education, geography, chemistry, agronomy, biology, and medicine (Moore et al., 2018; Bergeron and Gordon, 2017; Weeden et al., 2017; Rossello et al., 2020). Nevertheless, bottlenecks remain in engineering, physics, mathematics, computer science, and economics where they are enrolled in lower numbers; even though, also there, their participation has grown substantially (Verdín et al., 2018; Moore et al., 2018; Bergeron and Gordon, 2017;

<sup>3</sup> <https://www.newtimes.co.rw/section/read/197797> [https://www.buffpost.com/entry/rwanda-a-success-story-of-women-empowerment\\_b\\_5a4f1d87e4b0ee59d41c09ad7guccounter=1&guce\\_referrer=aHR0cHM6Ly91bi53aWtpcGVkaWEub3JnL3dpZ2VvR2VuZGVyX2ZVxdVFaXR5X2luX1J3YW5kYQ&guce\\_referrer\\_sig=AQAAAJnG0YmW6KazPs1ADUvc2NSpB59KuC47HVAMGvtX2Cyo0Qw\\_xS35b7Bu0jhYHQme9BKdI1OnQweKu3S0uijiNwoudF2jjifxskQ9qIMwL3sitV4d-cpEyEMt6tdbtmZEa5RB4fmL6cj-e4o8ivV\\_hftrf03B51fn3jMasYzj2g0](https://www.buffpost.com/entry/rwanda-a-success-story-of-women-empowerment_b_5a4f1d87e4b0ee59d41c09ad7guccounter=1&guce_referrer=aHR0cHM6Ly91bi53aWtpcGVkaWEub3JnL3dpZ2VvR2VuZGVyX2ZVxdVFaXR5X2luX1J3YW5kYQ&guce_referrer_sig=AQAAAJnG0YmW6KazPs1ADUvc2NSpB59KuC47HVAMGvtX2Cyo0Qw_xS35b7Bu0jhYHQme9BKdI1OnQweKu3S0uijiNwoudF2jjifxskQ9qIMwL3sitV4d-cpEyEMt6tdbtmZEa5RB4fmL6cj-e4o8ivV_hftrf03B51fn3jMasYzj2g0). Last access July 2020.

Weeden et al., 2017). Besides these significant changes and the increasing demand for skilled labour (Blair and Deming, 2020) women and people of colour have lower returns from their education (Saleh et al., 2017; Allen, 2018) and are less likely to find a job that matches their skills (Somers et al., 2019; Hunt, 2016).

This situation is particularly problematic in top positions in knowledge-intensive occupations, where females and blacks are severely under-represented. Among 500 corporate boards seats more than 66% are still held by white men;<sup>4</sup> and more striking, among the Fortune 500 CEOs 93% are males<sup>5</sup>. In 2019, only 29 out of 195 countries have a woman serving as heads of state or government.<sup>6</sup> The European Parliament (elected in 2019) has 41% of women and is the most gender-balanced ever elected.<sup>7</sup> However, many argue that with Brexit ethnic minorities will be even less represented in the EU Parliament.<sup>8</sup> 22% of the European population is black, but only 3 of the 751 EU parliamentarians are black.<sup>9</sup> Similar issues are problematic also at universities. In 2016, less than 24% of European full professors were females.<sup>10</sup> Also the type of professor contracts and stipend varies with gender and race. In 2016, in the US, only 27% of full-time professors are female and 4% are black.<sup>11</sup>

The previous examples highlight that the main issue remains a lack of transformation at the top of the hierarchy of knowledge-intensive occupations.<sup>12</sup> Given the increasing number of qualified women and people of colour, the lack of social transformation in top positions is, in many cases, hard to explain without considering network mechanisms. Network mechanisms are essential to successful entry into many knowledge-intensive occupations. To a very great extent, the fundamental process by which transformation happens is the entry one, following by subsequent promotions. In those jobs, fundamental tasks of the entry process are training and supervision, collaboration and mutual evaluation. Those tasks imply that the work is often organised in teams and, therefore, those workers happen to know each other. In these circum-

<sup>4</sup>Available at <https://www.catalyst.org/research/missing-pieces-report-the-2018-board-diversity-census-of-women-and-minorities-on-fortune-500-boards/>; access March 2019

<sup>5</sup>See Barnard et al. (2016)

<sup>6</sup>[https://en.wikipedia.org/wiki/List\\_of\\_elected\\_and\\_appointed\\_female\\_heads\\_of\\_state\\_and\\_government](https://en.wikipedia.org/wiki/List_of_elected_and_appointed_female_heads_of_state_and_government); last access July 2020

<sup>7</sup>Available at <https://europarl.europa.eu/election-results-2019/en/mep-gender-balance/2019-2024/>; access March 2019

<sup>8</sup><https://www.independent.co.uk/news/uk/politics/brexit-bame-eu-parliament-members-ethnic-minority-a9315036.html>; last access July 2020

<sup>9</sup>In 2014 elections <https://www.theguardian.com/world/2018/aug/29/eu-is-too-white-brexit-likely-to-make-it-worse>; access March 2019

<sup>10</sup>see <https://www.catalyst.org/research/women-in-academia/>; last access July 2020

<sup>11</sup><https://nces.ed.gov/fastfacts/display.asp?id=61>; access March 2019

<sup>12</sup>The phenomenon is often defined as the glass ceiling effect (Lewellyn and Muller-Kahle, 2019).

stances, social networks not only emerge, but they are fundamental in characterising the profession in its entry (and promotion) mechanism.

When social networks emerge, socio-economic theory highlights that mechanisms such as homophily, Matthew effects and prestige stratification influence how a system works and resources are allocated. However, the mentioned social network mechanisms have largely been theorised in developed contexts. In addition, most of the available empirical results focus on the U.S. or European countries. The lack of empirical evidence in “less developed” settings would not be problematic except that systemic mechanisms may differ in contexts where resources are scarcer or allocated differently. Emerging economies have less mature, smaller and expanding knowledge sectors and, contextually, they are often characterised by different social norms. In similar situations, the distribution of resources and norms may interact with network mechanisms in the workplace. The interaction of thereof could potentially amplify the extent of network mechanisms in the entry process. However, the specific context of developing knowledge systems may also attenuate those mechanisms. For example, on the one hand, a social norm that blindly attributes “wisdom” to senior members could reinforce Matthew effects – a young member of the organisation will employ more time to gain independence from a senior one. In contrast, on the other hand, this social norm and the strength of Matthew effects could be counterbalanced by the expansion and dynamism of knowledge-intensive sectors often experienced by emerging economies. When a knowledge sector is growing, experienced professionals are often scarce and it might be easier to have social transformation because organisations will have incentives to hire younger agents. Where the letters are both more likely to have higher levels of education and to belong to previously excluded groups. Therefore, the study of systemic mechanisms in contexts where knowledge-intensive workplaces are emerging may offer new insights about their functioning and motivates the empirical investigation outside North America or Europe.

Given those considerations, this thesis investigates the entry process in the workplace and discusses its relation with social transformations. The ultimate objective is to uncover the mechanisms of the entry process and whether they can inhibit/foster social transformations. In other words, what are the processes that affect the differential entry rate of under-represented groups<sup>13</sup> (e.g. female, people of colour) into desirable

---

<sup>13</sup>Related to under-represented groups today is popular in the U.S. the definition of Under-Represented Minorities (URM). The term refers to African American, Hispanic, Native American and Mixed Race. However, the term is hardly applicable outside the U.S. context. In many countries, the groups which are under-represented in the workplace are the majority in the general population. For example, In South Africa, 90% of the general population is black. The black population was segregated during apartheid and

(competitive and high skilled) occupations? To answer this elusive question, I will consider the specifics of a competitive and knowledge-intensive profession.

Beyond the extensive body of literature that has investigated similar topics in the past, the focus of this work is rather specific. I will focus my discussion looking at the process of entering into academia in South Africa and Mexico. In contrast to most of the literature, I will consider three specific phases preceding an academic occupation.

Academics have their work organised around three pillars: supervision (or training), collaboration, and mutual evaluation. Thus, I will explore the social network mechanisms of homophily, Matthew effects, and prestige stratification where supervision, collaboration, and mutual evaluation first manifest in the career. In particular, I will examine three phases of the entry process. The first phase that I will look at is the student-supervisor relationship in the educational phase. Supervision during the educational phase influences students' future development. In this phase, students form their beliefs about the world and generate perspectives on their future careers. Moreover, the student-supervisor relation may provide role models for students. Specifically, I will look at the role of race and gender homophily in the formation of the student-advisor relation in the educational phase.

Further, as the second phase, I will study co-authorship during the Ph.D. This phase is critical since nowadays, obtaining a doctoral degree is a necessary condition to become a scholar. Moreover, doctoral (publishing) productivity is often an important criterion for screening Ph.D. job market candidates. In particular, I will analyse whether homophily and Matthew effects affect students' productivity during their doctorates.

Lastly, as the third phase of the entry process, I will consider the transition into the first academic job where the mutual evaluation of quality between job applicants and the hiring committees implies an emerging prestige ordering among institutions

<sup>14</sup> The transition into the first academic job after the Ph.D. represents the first step into an academic profession. I will use the mutual quality evaluation to derive the hierarchical stratification of the universities by prestige and ask how the transition in this hierarchy affects scholar performance.

The following sections of the chapter elaborate on the general framework of the thesis.

---

is still under-represented in top job positions. Similarly, in Mexico, income inequalities are severe. People of Indigenous descent are 60% of the populations but under-represented in top positions and at higher risk of poverty. See <https://theconversation.com/study-reveals-racial-inequality-in-mexico-disproving-its-race-blind-rhetoric-87661> I will generally refer to under-represented groups where the word under-represented refers to the workplace. This allows me to include in the general discussion broadly all discriminated or disadvantaged groups. The advantage of the term is that is broad and includes female, people of colour, ethnic minorities, people with disabilities, migrants, sexual minorities.

<sup>14</sup>The job applicant applies after she/he accesses the "quality" of a faculty and the hiring committee search for the "best" candidate. See chapters 4 and 5.

## 1.1 Entering process

The entry process into a profession comprises several phases, such as training, job searching, creation of collaborations, that culminates with a hiring decision. After hiring, often, the entry process is not exhausted and starts a phase of evaluation. During the phases pre- and post-hiring, the newcomer has to create linkages with co-workers and builds her/his reputation. I will argue that in knowledge-intensive jobs, pre- and post-hiring phases are as crucial as hiring decisions. In contrast with most occupations, supervision (or training), collaboration, and mutual evaluation are three essential job-tasks that differentiate knowledge-intensive jobs from others.

However, often the empirical research disregards the characteristics of knowledge-intensive occupations and concentrates in a hiring (or promotion) decision in isolation. Specifically, past research overlooks the fact that hiring decisions, do not stand in isolation, but are part of a series of phases that compose the entry process into a profession. Indeed, in any hiring decision, an applicant's quality and productivity obviously play an important role, but beyond those two considerations, past research underlines that to enter into a profession successfully, status and social connections are fundamental (Bidwell, 2011; Brink and Benschop, 2014; Zinovyeva and Bagues, 2015; Burris, 2004; Han and Han, 2009). For example, numerous multinational organisations before job interviews exchange information about job candidates, and exchange managers and engineers between each-others (Westphal and Zhu, 2019; Edström and Galbraith, 1977; Han and Han, 2009). Indeed, it must be acknowledged that status and social connections are part of a social capital which develops in the pre- and post-hiring phases.

Many top job positions in knowledge-intensive sectors require a training period (often pre-hiring) which is supervised by a senior member. More generally, the phases of the entry process into a profession involve not only the job candidate, but also supervisors, co-workers, evaluation committees, and organisational/institutional characteristics. In those phases, status and social connections can potentially bias hiring decisions. This may happen because often the same set of agents are involved and called to tutoring, teach, evaluate and decide hiring.

### 1.1.1 Network (or Systemic) mechanisms of the entry process

The involvement of many agents in the entry process entails that social networks and their mechanisms naturally emerge. Thus, in order to examine the entry process, the study of network mechanisms that operate at the individual and aggregate level is fundamental. Those mechanisms are not only crucial for the understanding of the

entry process in general but, are essential to understand what are the constraints that under-represented groups face in entering into desirable occupations. In these respect, sociology identifies homophily, Matthew effect, and prestige stratification as the main systemic mechanisms creating social inequalities.

Indeed, a network mechanisms as homophily, for example, may favour the majority group and disadvantage others. Further, it must be acknowledged that most of network mechanisms operate at different levels of aggregation. Thus, their effects could be different at the individual, group, and system level. For example, homophily, as described below, could be beneficial for an individual but at the same time may reduce the total number of relevant connections between one group and another creating some degree of segregation that could negatively impact the aggregate functioning of a system. With these premises, I will focus on homophily, Matthew effect, and prestige stratification in the entry process.

## **Homophily**

Homophily (or assortativity) refers to the tendency for agents to connect disproportionately to those similar to themselves. The similarity of characteristics can refer to exogenous (e.g. gender, ethnicity, age, nationality) or endogenous (e.g. income, social status, values, norms) factors (Kossinets and Watts, 2009; Pin and Rogers, 2016). Exogenous characteristics (often) correlate with endogenous ones. The latter complicates the attempt to study the different mechanisms through which homophily arises.

The mechanisms driving homophily originate both at the individual level and the aggregate level (Smith et al., 2014; Granovetter, 1973; Barnett and Benefield, 2017; McPherson et al., 2001; Currarini et al., 2009; Podolny, 1994; Jackson, 2005). At the individual level, it arises when agents are endowed with preferences of forming connections with those of the same type. At the aggregate level, it arises when the potential meeting pool of agents is homogeneous along with some characteristic but not representative of the population; so, even when agents form ties at random (type-blind), they are more likely to form same-type of ties, irrespective of their preferences. Both mechanisms of homophily formation influence how ties arise in social networks. Therefore, irrespective of its source, homophily produces a recognisable structure in networks: tightly connected communities that appear to be self-organised around some common characteristics/traits which are observable at the aggregate level. In other words, homophily creates some degree of segregation in the network across groups of people with different traits.



In terms of aggregates, a segregated structure influences the total amount of social capital of each group. For example, the less represented (or peripheral) group will have more constraints, as a group, in forming same-type of links, and thus, lower possibilities to create network effects and a lower social capital. Therefore, homophily is potentially harmful at the aggregate level because it reduced the number of relevant connection among different groups.

Studying the individual, a large body of literature underlines that homophily in an agent's social network generates social capital which is essential for finding a job (Granovetter, 1973), hiring a new co-worker (Fernandez et al., 2000), training (Hezlett and Gibson, 2007), and referrals (Zeltzer, 2019). Therefore for individuals, homophily can be beneficial. The benefit of homophily can arise in specific contexts. For example, newcomers have few connections and often lack of experience. In such a case, homophily may help an agent to overcome these handicaps. Indeed, empirical results confirm that newcomers tend to display higher levels of homophily in their local networks (Mollica et al., 2003; Pillai et al., 2018).

In an entry process, the homophilic tendency of individuals could be a way to make decisions of association when information is imperfect. Related to this hypothesis, Kossinets and Watts (2009) find that homophily governs new tie formation when agents enter a new social locus; that is when they have no friends or friends-of-friends in common with incumbent members of the group they want to enter. This observation is crucial to any entry process which often involves entering into a new group (or club) of people and making new connections. Therefore, homophily may represent a general mechanism to reduce the risk of entering a new "club" commonly used by agents. Indeed, its presence is documented extensively (Smith et al., 2014) in social (McPherson et al., 2001; Currarini et al., 2009), economic (Podolny, 1994; Jackson, 2005), and cultural (Barnett and Benefield, 2017) networks. However, such a ubiquitous mechanism implies that under-represented groups may encounter significantly more entry barriers than others.

The barriers expressed by homophily in tie formation within an organisation can affect, especially in the entry process, socialisation, the acquisition of information or skills, and performance of the newcomers.

The reasons people socialise more with those with similar traits can be numerous. It may simply be because communication is easier, more comfortable, or because individuals are used to it due to structural constraints. Beyond the motives of homophily,

its presence affects how people socialised and is relevant in terms of information or skills acquisition. In terms of skills, a close vis-à-vis relation is often conducive for acquiring tacit knowledge while anecdotes or small-talk are a useful source of information. Concerning homophily and performance, some management studies find a positive association (Frink et al., 2003; Watson, 2012; Gompers and Wang, 2017; Fang and Huang, 2017; Ertug et al., 2018; Kilduff and Lee, 2020).<sup>15</sup> In particular, Ertug et al. (2018) find that homophily is positively related to performance, but only for individuals of low job rank, who are more likely to be newcomers in the organisation. Thus, homophily could be seen as a mechanism which helps newcomers to overcome their disadvantage with respect to incumbents in organisations.

### Matthew effects

The Matthew effect, cumulative advantage, or rich-get-richer effect refer to the process where some nodes (agents), which are endowed with more links (resources, or performance) than others, tend to accumulate more and more, becoming disproportionately more influential (rich, or productive) than others Zuckerman (1977); Barabási and Albert (1999). This process is observable when network effects are at work. The Matthew effect also manifests at the aggregate and individual level.

At the individual level, this process of cumulative advantage generates agents that act as hubs where their initial success generate later recognition and prestige. So their statuses are sometimes irrespective of their intrinsic “quality”. Robert Merton, in his famous paper in *Science*, wrote with respect to emergent hub scientists:

«[...] Eminent scientists get disproportionately great credit for their contributions to science while relatively unknown scientists tend to get disproportionately little credit for comparable contributions [...] The world is peculiar in this matter of how it gives credit. It tends to give credit to already famous people [...] As we examine the experiences reported by eminent scientists we find that this pattern of recognition, skewed in favor of the established scientist, appears principally (i) in cases of collaborations and (ii) in cases of independent multiple discoveries made by scientists of distinctly different rank.» — pag. 57 — Merton, Robert K. "The Matthew effect in science: The reward and communication systems of science are considered." *Science* 159.3810 (1968): 56-63.<sup>16</sup>

<sup>15</sup>Often results vary depending on the context and the nature of decisions or task that are involved.

<sup>16</sup>Indeed, Merton's wife Prof. Harriet Zuckerman conducted the interviews reported in the paper and was later credited by Merton as the co-author of the paper. However, she was an invisible collaborator and her work did not receive the same credit (Rossiter, 1993).

The Matthew effect in the entry process implies that some agents start in an advantageous position (or are significantly differently equipped at the beginning) than others. Therefore, the patterns that those agents have in the entry process are different with respect to the rest of the population. At the individual level, for example, agents entering the same job may have different experiences and connections with co-workers. In this respect, Bidwell (2011) compares individuals who experienced different entry processes in a financial services company in the US. Those promoted internally and those hired externally. He finds that those internally promoted have better performance for the first two years and lower exit rates compared to those hired externally. A similar finding indicates how the accumulated experience/knowledge of the workplace generates a Matthew effect that affects performance.

### **Prestige Stratification**

Prestige stratification is the process that hierarchically differentiates groups or organisations according to the level of respect/consideration they have in a community.

Prestige stratification is problematic when quality is difficult to measure and reward mechanisms are imperfect. In such cases, prestige stratification creates a hierarchy inherited from the past which may confer unmotivated advantages (Burris, 2004). In particular, when prestige stratification is coupled with the Matthew effect small initial prestige differences can be amplified, massively perpetrating inequalities irrespective of individual merit.

Concerning the entry process, prestige stratification is linked, in particular, with hiring mechanisms. This because hiring involves a mutual evaluation of “quality”. Both the job applicant and the hiring committee evaluate each other’s “quality” to make their decision. For example, empirical work finds that agents who studied or worked in a prestigious institutions are perceived as more competent than others who do not and are in an advantageous position in terms of hiring decisions, and promotions (Araki et al., 2016; Jung and Lee, 2016).

As summarised above, homophily, Matthew effect and prestige stratification may bias the entry process to the disadvantage of previously excluded groups. The three mechanisms, thus, may constrain social transformation. In what follow, I will focus on how homophily, Matthew effect and prestige stratification act in an academic

workplace.

## 1.2 Entry in a job in academia

Although data on the entry process in the private sector are often not available or limited, similar data are available for academia. Academia is a highly competitive and knowledge-intensive profession where the entry process and social transformation follows similar mechanisms to other sectors. Indeed, teamwork, individual output, social connections, and training are determinants of the entry process also in academia.

Academia matters a lot to society and the economy of a country. Universities are part of the innovation system of a country and they are foremost in knowledge production and diffusion. Scholars create and direct new knowledge and scientific paradigms (Kuhn, 2012). The knowledge created in universities contributes to the development of new technologies and innovations. The modern knowledge economy highlights the importance of a country to absorb new technologies. In this, the higher education system has a primary role because it is responsible for the culture, education, and development of future professionals. Therefore, the academic profession has essential responsibilities to society and biases in the entry process there, might distort the quality and quantity of the knowledge created.

Luckily, information related to scholars is available in all phases of the entry process often also for developing contexts. I will take advantage thereof and focus on academia in the empirical investigation of this thesis. In this section, I will discuss the traits of entering academia.

The main phases characterising the entry process in academia are the educational phase (bachelor, master), the training phase (Ph.D.), and the first academic job (post-doc, researcher).

The educational phase comprises mainly bachelor and master level.<sup>17</sup> University studies open doors to future careers but at the same time conditions people about what sorts of careers are relevant for them (Gersick et al., 2000; Wenger, 2010). In this respect, the master level is particularly formative and crucial for the decision to apply to doctoral programs. The kinds of relationships people have in university can be formative in the way they view the world and what they expect as a “natural” relationship. For Astin and Astin (2000) university classrooms enable students not only to acquire knowledge and skills but also to develop personal qualities and the

---

<sup>17</sup>In some cases, doctoral programs offer courses.

ability to work collaboratively. This development arises both through formal learning and from interactions with colleagues, peers and teachers. For many degrees, the final requirement is a thesis, and this is typically supervised or mentored by a faculty member, who guides the student not only in her/his dissertation but often about much wider subjects such as career prospects, future education and so on, as well.

The training phase is the most critical to enter academia. Training in academia corresponds to doctoral education and is supervised by a professor. Obtaining a Ph.D. diploma is nowadays necessary to enter academia also in developing countries. Moreover, the doctoral training represents the first step of a career and often the first co-authorship. During the Ph.D., students acquire the skills needed for the profession. It is there that they acquire the ability to conduct independent and publishable research, to work collaboratively, to evaluate each-other's contributions, and to teach and tutor. The role of social connections or social networks is crucial also in academia, and particularly so for young scholars. So creating connections is an essential part of an early career. Like so in many social contacts, the process of forming connections in academia may be subject to homophily. Additionally, from the student's point of view, properties of supervisor and institutions matter a lot since they can serve to enhance network effects (e.g. Matthew effect, prestige stratification). For example, a student can acquire co-authors from her/his supervisor's network. Moreover, in many disciplines, also the positioning and the quality of connections within a laboratory (or department) highly influence "quality", performance, and future hiring perspectives (Brink and Benschop, 2014; Zinovyeva and Bagues, 2015; Ramasco and Morris, 2006; Ma et al., 2020). This may imply that the acquisition of skills and the likelihood of entering academia depend not only on the quality and effort of the student in isolation but also on how she/he is able to accumulate social capital.

After the doctoral degree, a future scholar enters the job market. The first academic job is a post-doc or a position as a researcher. The Ph.D. job market comprises job market candidates and university institutions. From the side of job market candidates, a "good" affiliation can enhance access to colleagues, collaborations, and resources. From the side of institutions participating in the Ph.D. job market to hire a promising researcher can improve their performance as well. Indeed, young researchers will turn into lecturers and professors that will stay in academia for a long time. However, for institutions, hiring decisions suffer from imperfect information because young candidates often have thin research records. In such situations, a university hiring committee will look for signals of quality, one of which is the Ph.D. granting university. In this way, doctorates from "famous" universities have an advantage in the market.

The behaviour of the agents involved in the Ph.D. job market, thus, creates a network with predictable hiring flows from a university to another that contains a stratify prestige hierarchy (Clauset et al., 2015; Burris, 2004; Cowan and Rossello, 2018; Gonzalez-Sauri and Rossello, 2019).

The structure of prestige stratification of the Ph.D. job market may imply profound inequalities. For example, individuals that for various reasons studied at institutions which are less “prestigious” (or integrated in hiring networks) are disadvantage in the job market and will hardly be hired in top-tier universities.

### 1.2.1 Homophily in tie formation

As described above, homophily is a fundamental mechanism of social interaction. Concerning its formation, homophily has two sources. One at the individual level and another at the aggregate level. In the literature the first is called choice homophily while the second is induced homophily. Choice homophily arises when individuals are selected into a position for conscious or unconscious preferences or tastes that can be coupled with imperfect information (see the taste for discrimination (Becker, 1957) and statistical discrimination hypothesis Phelps (1972)). In contrast, when homophily that has its source in the aggregate level is independent of individuals and induced by the surroundings. It arises when the potential meeting pool for agents is homogeneous along with some characteristics because of network structures, path dependence and segregation.

In Higher Education, the evidence of induced homophily is observable (mainly) in the demographic composition across the academic ranks and fields. At high job-ranks and in prestigious institutions, white males still dominate academic positions. For example, in 2016, according to the U.S. Department of Education, only 27% of full-time professors are female, and 4% are black.<sup>18</sup> Across fields, a certain degree of segregation is evident where female and black faculty concentrate in specific disciplines (e.g. art, history, biology, medicine) and are under-represented in others (e.g. engineering, physics, computer science, and economics) (Moore et al., 2018; Weeden et al., 2017). Of course, part of the observed homophily here can be driven by choices of individuals too.

The answer to a key question to understand demographic imbalances at universities is therefore elusive: Is there a bias in selection committees, or is the population of applicants not representative of the population as a whole?<sup>19</sup> To answer this question,

<sup>18</sup>Source NCES <https://nces.ed.gov/fastfacts/display.asp?id=61>; access March 2019

<sup>19</sup>UC Berkeley was accused of (anti-female) gender bias in undergraduate admissions in the 1970s, but on closer examination, it was uncovered that most of the discrepancy in admissions rates for male and

we need to understand one of the most fundamental systemic mechanism of social interaction: homophily. And in particular to understand the relative magnitude, in producing these imbalances, of homophily that arises at the individual and aggregate level. In other words how much of the homophily is induced and how much is of choices?

At both levels, empirical research has found evidence of homophily in academia. De Paola et al. (2015) studies the effect of homophily at the aggregate level. They find, controlling for various individual characteristics including “quality”, that women are less likely to apply to Italian Scientific Qualification competition in fields where females are under-represented, or in other words, where the level of induced homophily is high. Looking homophily at the individual level, Bagues et al. (2017) find, examining promotions to associate and full professor in Spain, that when there is a woman on a committee, the men on that committee become homophilous, and thus, less favourable towards female candidates.

Beyond committee decisions, scientific productivity can also be affected by homophily. Indeed, publications are often necessary for promotion and precondition for research funds. In this respect, Helmer et al. (2017) find homophily to be present in the peer-review process, finding that editors (male and female) are more likely to publish papers from colleagues of their same gender. Thus, since women are under-represented in many editorial boards, homophily may significantly reduce the number of papers published by female authors.

Homophily in higher education seems to be widespread within the different processes in which an academic is involved, such as credit from teamwork (Sarsons et al., 2015), promotion (Zinovyeva and Bagues, 2010), publishing (Pezzoni et al., 2016; Helmer et al., 2017; Rossello et al., 2020), and grant applications (Van Der Lee and Ellemers, 2015). One specific process that may matter a lot, but is as yet very little studied, is the selection of supervision. The student-supervisor relation is relevant in terms of education and affects future student career. Indeed, the doctoral advisor is often the first co-author of a future scholar. I will concentrate, thus, the study of homophily in the formation of the student-advisor ties.

In more general terms, the education literature finds that, at universities, an effective student-supervisor relation includes: (1) a focus on achievements and knowledge acquisition; (2) support, direct assistance, and role modelling; (3) mutual benefit; (4)

---

female applicants was induced by differences in the departments to which the groups applied. In fact, there was evidence that choice homophily operated in favour of admitting women (Bickel et al., 1975).

a personal nature beyond strictly work-related issues; and (5) a professor having experience, influence, and achievement (Jacobi, 1991; Gersick et al., 2000; Girves et al., 2005; Crisp and Cruz, 2009). These attributes involve social relations that are subject to homophily. Furthermore, when these attributes are present in a positive sense they can reduce attrition rates, and are found to be crucial, in particular, for under-represented and minority groups in early career stages (Girves et al., 2005; Terrell and Hassell, 1994). Indeed, role models are crucial, and mentors of the same race/gender of the students are often more effective (Bettinger and Long, 2005; Lockwood, 2006).

Thus, the student-supervisor relation, mainly for its relevance for under-represented groups, is crucial for social transformations. Effective supervision may help formerly excluded groups to enter organisations where they are under-represented and could be a channel for inclusion. However, student-supervisor is also a social relation where homophily can affect tie formation. Homophily in the formation of student-supervisor ties, because many traits (number of co-authors, productivity, rank, awards, sub-field) often correlate with gender and race, can influence the social capital of future academics. Besides, the presence of homophily in tie formation can make the entry process of under-represented groups harder because it limits their interaction with the former majority groups.

### **1.2.2 Doctoral productivity**

Connections are a source of social capital. But it is clear that the type and quality of connections and the characteristics of the agents involved matters for the nature and strength of social capital. Especially at early stages of a career, in forming connections homophily and the Matthew effect can play a role. Homophilous connections have been observed to be important for students' role models (Breda et al., 2018; Bettinger and Long, 2005; Lockwood, 2006). In addition, homophily can influence individual performance. For this, homophily can be relevant for doctoral productivity. The literature on education underlines that role models are important for students performance and that race and gender play a role (Girves et al., 2005; Terrell and Hassell, 1994; Main, 2014; Bettinger and Long, 2005; Lockwood, 2006; Aguinis et al., 2018; Breda et al., 2018; Rossello et al., 2020).

This literature finds that on average same-race/-gender supervision are more effective than cross-race/-gender. Main (2014) studies gender homophily using data from the U.S. Graduate Education Survey. In her sample of 7834 Ph.D.s in Social Sciences and Humanities from cohorts in 1982-1993 she finds that female students are more



likely to choose a female advisor and when they do so they are more likely to finish the Ph.D. and to complete it in a shorter period. In subsequent work, Pezzoni et al. (2016) study student-advisor gender couples differences in doctoral publication productivity. They examine 6 Ph.D. cohorts (2004-2009) from the California Institute of Technology and find no gender difference in productivity between same-gender couples. However, looking at cross-gender couples, they find that women working with male advisors publish 8.5% less than male-male couples.

Experiments, also, highlight a positive effect of same-gender role models. For example, in a recent intervention in France Breda et al. (2018) randomly assign high school classes to a lesson with a female scientist. They find that a female scientist role model increase by 38% the share of female students subsequently enrolled in selective science programs.

Training and research output are crucial to enter academia and, as documented in previous research, they are influenced by homophily. Indeed, a strong positive association between homophily and performance may imply that lack of social transformations at the top of the hierarchy inhibits transformations at the bottom. For example, the scarcity of black (female) full professors that can supervise Ph.D. students can have a negative impact on the average productivity of black (female) students. A similar mechanism, will thus decrease the likelihood that black (female) students enter academia and become future scholars, creating a vicious cycle.

Disparities across academic ranks has been studied in past research showing that female (Zinovyeva and Bagues, 2010; Mairesse and Pezzoni, 2015) and black (Fang et al., 2000) faculty are promoted at lower rates than white males. Looking at ethnic minorities (Fang et al., 2000) find this evidence for US medical scholars (N=50145). While, focusing on gender, Mairesse and Pezzoni (2015) analyse 1813 French physic faculties and find that a 40 years old female scholar with an average productivity (and quality) has 16.3% lower chances of being promoted than an equivalent male colleague. This may imply that Ph.D. graduates from under-represented groups gain scientific “independence” after a longer time compared to white male. Moreover, an high academic rank is important for founding, status, prestige and to attract students and collaborators. Thus, a similar promotion bias may reduce the chance for blacks and females of enhancing Matthew effects in productivity also later in their careers.

### 1.2.3 First job

Following the educational and training phase, a future scholar enters the academic job market. The first academic job is fundamental to enter academia successfully. The “quality” and size of the university of the first hire, the interaction with colleagues, facilities, students can enhance or inhibit the Matthew effect and so promotion and future career perspectives.

In order to navigate the academic job market, agents need to understand its dynamics. However, for newcomers, the academic job market is often cryptic. Past research underlines that many Ph.D.s do not understand the dynamics of the academic job market (Fox and Stephan, 2001; Akerlind, 2009; Felisberti and Sear, 2014; Hayter and Parker, 2019). And, many doctoral students leave academia when they realize the difficulties associated with entering it (Hayter and Parker, 2019; Stephan and Ma, 2005; Golde, 2005). The latter is potentially more problematic to previously excluded groups that for the mechanism of homophily (there are few professors “like” them) may lack mentorship.

The specific characteristics of the academic job market are a small turnover, hiring constraints, and prestige stratification of institutions. In relation to the size of the job market, Hayter and Parker (2019) underline that less than the 15% of post-docs in the U.S. will find employment as tenure-track faculty. Universities have hiring constraints which relate with the number of students and courses, number of retired professors, private and public investments, and sizes. These characteristics are included in formal university rankings that aim at evaluating the prestige of institutions.

Beyond the way in which university prestige is measured, it creates a structured and stratified university system. In this respect, the sociological literature underlines that predictable hiring flows between universities emerge. These flows follow a hierarchical and stratified hiring network where university prestige plays an important role (Webster et al., 1991; Burris, 2004; Jung and Lee, 2016). University prestige is the reputation that an institution has in a community of practice and is crucial in the transition in the first academic job. Scholars at the beginning of their careers can give few signals of their quality as researchers. For example, they have few records of publications and co-authors, and their specific contribution is often not observable. For doctoral graduates, the absence of signals of “quality”, makes the prestige of the university granting Ph.D. a signal for it. Thus, a prestigious Ph.D. increases the likelihood of being hired in a “good” first-job institution and this confers an advantage that

can enhance Matthew effects and future productivity (Mai et al., 2015; Burris, 2004).

## **1.3 Social Transformations and the Entry Process**

Social transformations are changes in the shared understanding of social norms in a society followed by a change in behaviour. Such a change is often lengthy because social norms are resilient. What constitutes their resilience is that norms come to existence because different groups share beliefs, and beliefs coordinate, codify, and sediment over time. In a similar manner, change is often a gradual phenomenon. In fact, the more common social transformations verify when little changes accumulating and stratifying over centuries. In this way, change gradually transforms a society. However, there are contexts or eras, where social, political, and economic conditions accelerate those changes.

Often, the labour market is one of the places where social transformations first become evident. This because often excluded groups are a source of new employees. Social transformations generate two observable outcomes: a change in the demographic composition, and a change in the structure of group interactions (Smith et al., 2014). Within a workplace, the social transformations change the pull of young job applicants which then affect the entry process and the composition of organisations. In other words, social transformations and the entry process are interlinked by the succession of job-cohorts, generations, hirings, and retirements. This turnover makes the relation between the new and the old generation of job applicants and senior members essential to effective social transformation. The interaction of the demographic groups in a population, the level of homophily, their hierarchy and the relation between the old and the new generation are particularly relevant in the entry process of newcomers. The entry process changes the demographic composition of a workplace and is an essential step for social transformations.

### **1.3.1 Mechanisms that foster or hinder social transformations**

Often social transformations concentrate in specific organisations, fields or tasks, and job-ranks. This concentration has to do with the relation between social transformations and systemic mechanisms of the entry process. Indeed, homophily in tie formation, Matthew effects in productivity, and prestige stratification of institutions can hinder or foster social transformation.

The first mechanism that play a role in social transformation is homophily in tie

formation. In this respect, a broad empirical evidence finds that newcomers in organisations tend to form homophilous relations (Bramoullé et al., 2012) and these are often associated with higher outcomes (Frink et al., 2003; Watson, 2012; Gompers and Wang, 2017; Fang and Huang, 2017; Ertug et al., 2018; Kilduff and Lee, 2020) (See section 1.1.1). Thus, under these conditions, homophily can affects transformation in two ways. One the one hand, homophily can hinder social transformation for relative group sizes and segregation. In term of sizes, a group that is under-represented in an organisation has lower chances to meet agents of the same group and to form homophilous ties. While in terms of segregation, often diversity is concentrated in the job hierarchy and top positions remain often dominated by white males. Thus, sizes and segregation imply a disadvantage for members of under-represented groups because they have a lower scope to generate social capital (as a group) by exploiting the mechanism of homophily. On the other hand, looking at the individual level homophily can foster social transformation thanks to the role model effect. When a certain number of agents from under-represented groups are present in an organisation, they may act as role models attracting and encouraging the entry process of newcomers from under-represented groups.

The second mechanism relevant for transformations are Matthew effects in productivity. They imply a process of path dependence where the initial level of resources highly influence future productivity. So, a lack of resources or social capital in early stages of a career might impede (or undermine) future outcomes. If homophily in tie formation is at work, it could inhibit the productivity of newcomers of under-represented groups because they lack role-models and supervision. Alternatively, if newcomers of the former majority groups are less likely to form ties with young members of under-represented groups this may inhibit the Matthew effect in the careers of the latter ones. In a similar case, Matthew effect in the entry phase and the lower extents of it for under-represented groups will inhibit social transformations. In contrast, under-represented groups may be specialised in specific fields, niches, or tasks. Specialisation could generate a cumulative advantage and enhances a Matthew effect that positively affects their productivity in those areas. In this case, forming homophilous relations coupled with Matthew effects may foster social transformations. However, this may be limited depending on the size, reputation, and impact of those tasks or topics of specialisations.

The third mechanism that may influence social transformation is the stratification of prestige. The reputation and impact of an institution emerge from the consideration it has among peers in a period. The mutual evaluation among members of a com-

munity generates a stratification of prestige. Related to social transformations, often prestigious institutions are found to have low levels of diversity. In many countries, under-represented groups often do not access quality education in prestigious schools. Given the importance that a prestigious degree has in signalling “quality” in the job market, groups that for various reasons are under-represented in those prestigious institutions encounter an additional barrier. Thus, the stratification of institutional prestige may constitute a mechanism who prevent social transformation.

## 1.4 The Context

This dissertation studies the entry process in a knowledge-intensive sector in two developing contexts: South Africa and Mexico. In contexts such as these, knowledge-intensive jobs are evolving fast and are relatively recent. The knowledge sectors in middle income countries have several properties that differentiate them from those in advanced economies. For example, knowledge-intensive occupations, while growing, and considerer "the future" still make a relatively small part of employment. Consequently, training for those occupations is often relatively recent: doctoral programmes became systematic only in the last two decades.

In general, the entry process in the workplace is understudied in developing context because of lack of data. However, the focus of a developing context could uncover more clearly the systemic mechanisms of an entry process because networks are smaller and less integrated internationally. Similar characteristics make the study of entry mechanisms easier, and less demanding in terms of data. For example, one can think of the country as a relatively closed system. In the context of academia, it is a good approximation to assume that only domestic universities participate to the Ph.D. job market.<sup>20</sup> Thus, this work treats the South African, and the Mexican academia as separated to the rest of the world. In addition, middle income and developing countries are characterised by smaller knowledge systems where resources are often more scarce and concentrated. Concentration of resources makes flows from one university to another more predictable and network mechanisms stronger.

### 1.4.1 South Africa

South Africa is an example of a country undergoing rapid social transformations where the social, political, and economic conditions have accelerated changes. Apartheid

---

<sup>20</sup>If a university attracts international students it is likely to be the same that attracts the most also nationally.

legally divided all walks of life on ethnic and racial lines, for more than eight decades.<sup>21</sup> The situation escalated in the late 80s, the national and international political pressure of that time accelerated the change of the society. The turmoil culminated in 1991 with the abolishment of apartheid. With the elections of 1994, formally all institutions became democratic and the integration process started.

Since the new elections of 1994, the country is de-segregating and transforming from a society legally divided on racial lines to a more integrated one. However, the transformation process is not complete, and in many places there remains a considerable gap between the current situation and aspirations. Higher education is still for only a few, and the upper secondary school is represents one of the bottlenecks. Unicef reports an attendance rate of upper secondary school in 2019 of 70% with a completion rate of 48.<sup>22</sup>

On the top of that, the higher education sector has its struggles. During apartheid, the university system was segregated and essentially a two-tier system. Black, Indian, and coloured institutions were specialised in technical education and teaching, significantly underfunded, and not in the condition to do any meaningful research (Herman, 2017). The system has begun a process of integration, but still social transformations are unevenly distributed and they present contradictions and bottlenecks.

The main contradiction remains the heritage of a two-tier system in many geographical and research areas. For example, besides the significant transformation of the student body where the percentage of black students has increased from 20% to 61% in 20 years (Rossello and Cowan, 2019),<sup>23</sup> higher education in many areas of the country remains the privilege of only a few. The university system is relatively small. In a population of almost 60 million people there are 24 public universities and around 4000 full-time professors.<sup>24</sup> Furthermore, limiting access to higher education, the costs play a significant role. The average annual tuition fee is high, between 30000 and 64000 ZAR (approximately 1500-3200 Euro) in a country where gdp per capita

---

<sup>21</sup>I refer here to the “Natives Land Act” of 1913, that limited land property to 7% for the black population. The act represents the first discriminative law. Formally, the term of apartheid was politically used for the first time in 1916 and was implemented in 1948.

<sup>22</sup>See unicef tables at <https://data.unicef.org/topic/education/secondary-education/>; last access September 2020.

<sup>23</sup>Formally, in South Africa there are four “racial” groups: black, white, Indian and coloured. The word “black” is sometimes used to refer to the aggregate of black, Indian and coloured. We use this meaning throughout this text.

<sup>24</sup>The datum refers to statistics of the year 2012 available at <https://africacheck.org/reports/how-many-professors-are-there-in-sa/>; last access March 2020

(PPP) is around 232000 ZAR. (12000 EUR)<sup>25</sup> In addition, many Ph.D. programs are not funded (especially in the Social Sciences and the Humanities) and the production of Ph.D.s and research is concentrated in few institutions (Cowan and Rossello, 2018). The latter contributes to the persistence of demographic imbalances in the professoriate. In 2014, black scholars constituted 44% of the total and remained concentrated in some fields and universities (Rossello and Cowan, 2019).

On the other hand, bottlenecks arise in the commitment of South African academia to reach an international role and position in terms of research output and impact. This concentrates resources in formerly white universities in specific subjects. In those universities and subjects, the enrolment of black (South Africans) and females is low, creating scarcity of supply of potential future professors of under-represented groups. The university system of the country plays a crucial role in Southern Africa and attracts many international students also from other African countries (Confraria and Godinho, 2015). Moreover, the South African university system has competitive advantages and specialisation in many fields where it is producing frontier research (e.g. astronomy, biology, anthropology, medicine, geology, agronomy). But prioritising certain parts of the system does not permit the development of an African based Social Sciences and Humanities and Base-Science. Investing in those underfunded sectors may constitute an important engine for employment, innovation, and development in the country. Unbalances are clear looking at percentages of graduates by fields. Overall in 2012 the country 20% of students graduated in STEM subjects<sup>26</sup>

Unbalances and inequalities in the South African university system provoked two large student protests in 2015 demanding racial transformations and equal access and opportunities.<sup>27</sup> <sup>28</sup> These protests represent a puzzle for higher education policy where many argue that exist a trade-off between inclusion and quality (based on international standards).

### 1.4.2 Mexico

Mexico is a vast country with a population of more than 128 million people. In contrast to South Africa, racial and cultural integration was promoted since the beginning of the 20th century. However the country is plagued by income inequalities and unequal opportunities that interest a large part of the Mexican population.

<sup>25</sup> Available at <https://businesstech.co.za/news/finance/293800/university-fees-2019-how-much-it-costs-to-study-in-south-africa/>; last access April 2020

<sup>26</sup> See <https://www.businessinsider.com/most-technological-countries-lag-behind-in-science-2015-12?IR=T>; last access June 2020

<sup>27</sup> For details see [https://en.wikipedia.org/wiki/Rhodes\\_Must\\_Fall](https://en.wikipedia.org/wiki/Rhodes_Must_Fall). Last access March 2020

<sup>28</sup> For details see <https://en.wikipedia.org/wiki/FeesMustFall>. Last access March 2020

Indeed, access to quality education is considered problematic in Mexico today. While literacy is quite high, and over 90% of children attend primary school, only 62% of them go to secondary school and only a quarter of them pass to higher education.<sup>29</sup> A reason behind the unequal access to education is often the lack of infrastructures throughout the rural schools.

In Mexico, there are 1250 institutions of Higher Education, including public universities, technological institutes, technological universities, private institutions, teacher training colleges, and other public institutions. Among those, universities are 213. However, 50% of the research and 58% of the students concentrates in only 45 public universities which are mostly located near Mexico City and in other large cities.<sup>30</sup> Concerning the distribution of academic fields, priority has been allocated to applied STEM subjects. In 2012, Mexico was the 8th country in the world for the percentage of graduates in STEM subjects where 25% of students graduated in those subjects (the U.S. has 16%, and the Netherlands has 15%)<sup>31</sup> The latter could contribute to the persistence of inequalities and the marginalization of rural communities. Indeed, according to Oyarzún et al. (2017), the Indigenous population in Mexico faces not only exclusion from higher education for the low levels of access linked to geographical and income inequalities. Their additional exclusion connects to the little acknowledgement of the distinctive cultural traditions of those communities within the university curricula (Franco and McCowan, 2020).

In both contexts, common goals refer to the expansion and consolidation of their fast growing of knowledge sectors in order to acquire a solid international position and partnership. These goals are accompanied with numerous challenges in addressing the high social inequalities that characterised both countries. In this, the higher education system in both countries may play an important role in mediating goals and challenges.

## 1.5 Knowledge Gap

The study of constraints that under-represented groups face in the labour market has been extensively studied in many social sciences, like sociology, economics, and psychology. However, few study how those constraints relate to network mechanisms

---

<sup>29</sup>See [https://en.wikipedia.org/wiki/Education\\_in\\_Mexico#Quality\\_of\\_education\\_in\\_Mexico](https://en.wikipedia.org/wiki/Education_in_Mexico#Quality_of_education_in_Mexico); last access June 2020

<sup>30</sup>See <https://education.stateuniversity.com/pages/984/Mexico-HIGHER-EDUCATION.html>; last access June 2020.

<sup>31</sup>See <https://www.businessinsider.com/most-technological-countries-lag-behind-in-science-2015-12?IR=T>; last access June 2020



and social transformations in the entry process. Indeed Smith et al. (2014), with the broad focus of the U.S. society, links the network mechanism of homophily and social transformations using the U.S. General Social Surveys 1985-2004. They also underline, with the following sentence, how little has been done in the understanding of those phenomena:

“Given homophily’s central importance, it is surprising that we have so little knowledge of whether this fundamental social fact has changed over time. [...] Most of the information we have about changes over significant periods of time in the network structure of our society come from either the exclusive, close ties of marriage (Kalmijn 1998) or the much weaker ties of co-employment (Tomaskovic-Devey 1993), co-residence (Massey and Denton 1993), or co-matriculation (Shrum, Cheek, and Hunter 1988).” – pag. 433-4 – Smith, Jeffrey A., Miller McPherson, and Lynn Smith-Lovin. Social distance in the United States: Sex, race, religion, age, and education homophily among confidants, 1985 to 2004. *American Sociological Review* 79.3 (2014): 432-456.

Social networks and their mechanisms represent the fabric of our society. Each of us is involved in many networks at the same time. In this respect, any entry process (in clubs, organisations, or workplaces) involves the interactions of the different networks in which we are involved. For example, the likelihood that a friend becomes a co-worker (or vice-versa) depends on how (and according to which traits) the friendship network and the co-worker network overlap. Moreover, if the friendship network forms according to racial homophily, this will imply that under-represented groups will be less likely to get into that workplace by exploiting information from their friendship networks. However, the study of these sorts of mechanisms requires a large amount of data. For this, the lack of longitudinal and representative data of the entry process represents a gap in the study of social network mechanisms. Indeed, there are many hypotheses and models on how network mechanisms work but little empirical evidence<sup>32</sup>. Indeed, network mechanisms are complex and ambiguous. For example, homophily, Matthew effects, and prestige stratification have both aggregate and individual effects. In section 1.3.1 I described some possible aggregate and individual effects that may foster or inhibit the entry process of under-represented groups. This ambiguity of the effect of network mechanisms implies that little is known on how they relate with social transformations.

---

<sup>32</sup>See Jackson (2010) for examples on network models

Social transformations change the demographic composition of a workplace and how groups interact. In the workplace, the social transformation process over time first changes the pool of candidates. After, thanks to the entry process and subsequent hirings, social transformations filtrate and the demographic composition of an organisation changes. However, a demographic balance into desirable (knowledge-intensive and competitive) occupations is particularly problematic in developing contexts. There, this type of occupation is still relatively scarce and often accessible only to a privileged minority. An additional problem is the lack of possibility to monitor social transformations as they tend to take place over a relatively long time. Many developing countries do not have longitudinal data that would allow one to track the changes either as they take place or in retrospect. In such contexts, a place to look at network mechanisms in the entry process is academia. First, academia is a knowledge-intensive workplace where scholars collaborate, know, train, and evaluate each-other. Second, data on academics are available also in developing countries (at least in principle). However, most of the research on how academics interact is focussed on Europe and North America and often concentrates in one organisation or field. In contrast with much of the literature, this dissertation studies the entry process of South African and Mexican future scholars. Moreover, using register data from their public agencies that evaluate research, I cover a representative sample of future scholars in both countries.

The study of transformation processes in developing contexts allows observing university systems in transition into a mature phase. Constraints in the entry process of under-represented groups arise both at the individual and aggregate level. During a transition phase of social transformation, those constraints may work differently. A less mature knowledge system, for example, will have less international collaboration and it will depend more on local linkages. Thus a focus in developing context is crucial to understand social transformations but often neglected in the literature.

The additional gap that this dissertation intends to cover is the distance between the literature that focus on homophily, Matthew effects and prestige stratification at the individual level and the one that studies them at the aggregate level. On the one hand, the economic literature emphasises the individual mechanisms of the entry process as individual performance, mobility, and role models, and homophily of choice (that often call racial-/gender-bias or preferences). On the other hand, the sociological literature focuses on aggregate (or systemic) mechanisms like group relations, the level of homophily induced by an organisation, the relations between institutions, and prestige stratification. Starting with these promises, this thesis connects the two liter-

atures. To connect the two, is not easy and requires together different methodologies and perspectives. For this, I developed new techniques that I describe in the following section.

## 1.6 Methods

Network mechanisms have individual and aggregate effects that possibly affect the entry process of individuals. Given their complexity, in each chapter, I employ and develop a methodology suited to study both.

**Chapter 2** develops a new methodology which separates choice homophily from that induced by the system. The method solves causal issues creating a null model of random and unbiased tie formation that controls for network structures, supervision capacity, and small sample bias. Further, I design a model with interacting populations and use it to estimate homophilous preferences using aggregate department-level data of student, supervisor, and tie composition. The methodology is promising for monitoring aggregate and group level homophily in the entry process. Further, the method can also be used with sparse register data.

**Chapter 3** uses an econometric technique. In particular, I employ a panel data analysis with a quantile regression approach. This to investigate the link between student-advisor same- and cross-gender (and race) couples and productivity. In this chapter, I investigate the relation between homophily and productivity in the different student productivity-profiles. This accounts for the possibility that homophily/heterophily (same-/cross-gender -race ties) works differently for (hub) scholars with extraordinary levels of productivity who experience Matthew effects.

**Chapter 4** develops a new technique to study prestige stratification of institutions. Based on hiring networks, the algorithm ranks universities according to their reputation among academics. Further, I develop a new quasi-experiment method to test the effect of movements in the prestige hierarchy on research quality. The method is based on exact matching and compares the performance of matched pairs of scholars with the same training (or job) but different prestige transitions.

**Chapter 5** modifies the ranking algorithm to rank university prestige. The modified version of the algorithm takes into account that universities may recruit in a cyclical manner. In this way, the university ranking is dynamic over time and accounts for changes in investment flows, modes, news. Then, I apply the matching methodology used in Chapter 4.

## 1.7 Thesis outline

This thesis investigates the entry process in academia focusing on network mechanisms that may constraint under-represented groups.

As mentioned above, the entry process changes the demographic composition of a workplace and is a precondition and the main channel to social transformations. In general, the entry process follows aggregate mechanisms at the system level and mechanisms at the individual level.

Systemic mechanisms influencing the entry process are the interaction of the demographic groups in a population, the level of homophily, hierarchy and links between institutions, and the relation between the old and the new generation.

At the individual level, important determinants of the entry process are individual output, social connections, training, characteristics of the student-advisor relation, and mobility.

**Chapter 2** analyses the systemic mechanisms governing student-supervision relations. In particular, the chapter studies homophily along racial and gender lines in the formation of student-advisor relations. In this chapter, first, I ask what the relative magnitude of induced and choice homophily in the system is. And further, where choice homophily originates in the interacting populations of black and white (male and female) student and advisors.

**Chapter 3** studies the individual mechanisms of student-advisor relations, asking whether homophily enhances student productivity or not. More specifically, it examines whether student-advisor gender and race couples matter for publication productivity of Ph.D. students and whether it changes depending on the student productivity-profile.

**Chapter 4** investigates systemic and individual mechanisms together. In particular, the chapter studies the link between mobility and hiring channels in the Ph.D. job market with institutional prestige and research quality in the short, medium, and long-run.

**Chapter 5** expands findings of Chapter 4 looking at prestige mobility and performance in a larger university system.

**Chapter 6** synthesises the key results of the thesis, highlights the elements of the larger significance of the study and lays the foundation for future research and gives general policy recommendations.



## Chapter 2

# Homophily in student-supervisor tie formation

This chapter investigates homophily<sup>1</sup> along racial and gender lines in the formation of student-supervisor tie. In the period of analysis (1973-2014) South African Academia was interested by profound social transformations and reforms of the higher education system. Profound social transformations potentially can change how demographic groups relate one other in terms of homophily. Thus, to study the relation between the two in the first part of the chapter I decompose observed homophily in its choice and induced components; and in the second part, I study where choice homophily originates in the interacting populations of students and advisors.

In the first part, I develop a new method to separate choice homophily from homophily induced by the system. The new technique uses two levels of aggregation of the data, system and department level. It creates for each level a null model of random and unbiased tie formation using repeated permutations. In this way it is possible to

---

<sup>0</sup>This chapter is based on Rossello and Cowan (2019), UNU-MERIT working paper series. I would like to thank the comments and suggestions of participants of the 2nd Workshop in Economics of Science and Innovation at BETA in Strasbourg; participants at EAEPE 2019 in Warsaw; participants at the seminar at Università Cattolica del Sacro Cuore in Milan; participants at the seminar at the BE-Lab at Harvard University. As well as those of Dr. Emanuele Colonnelli, Dr. Daniele Giachini, Prof. Alan Kirman, Dr. Mindel van de Laar, and Tatenda Zinyemba. Financial support was provided through the Institut Universitaire de France.

<sup>1</sup>The use of the word homophily in this chapter should not be confounded with the concepts of sexism and racism. Of course, there are some degrees of overlapping as agents that are sexist and racist will display an homophilistic tendency in their behavior. However, sexism and racism is just the tip of the iceberg as also progressive and anti-discrimination agents are often found to display homophilistic tendencies.

control for small sample bias, supervisor capacity, and other confounding embedded in the network. The findings highlight that racial and gender homophily interest the formation of student-supervisor relations. The decomposition of homophily in its two components shows that roughly half of the homophily observed is induced by the composition of the departments and displays little changes over time. Overall, comparing choice homophily along racial and gender lines their magnitudes are similar.

In the second part, I investigate the origin of homophily of choice in the demographic groups of students and professors. To do that, I develop a model with interacting student and supervisor populations. Where in each of the two populations dichotomous types represent race or gender. With the use of aggregate department level data of student, supervisor, and tie composition, I estimate homophilous preferences using the model. This new methodology offers a promising way to monitor, in the presence of sparse register data, social transformations within organizations. The results underline a high tendency to form same-type of tie among white (male) students and black (female) professors. Further, the model underlines that under some circumstances homophily can mechanically rise. During a transition phase of integration the changes in the relative sizes in the sub-populations can cause a mechanical increase of homophily, without changes in the behaviour of association among groups.

## 2.1 Introduction

“Homophily” refers to the tendency for like to be associated with like, and is often manifest when a particular social milieu is dominated by one group. Today we observe many contexts in which the presence of homophily is taken to be problematic: 66% of Fortune 500 corporate board seats are still held by white men;<sup>2</sup> and 93% of Fortune 500 CEOs are male<sup>3</sup>; only 37% of the European Parliament (elected in 2014) is composed by women;<sup>4</sup> of the 751 EU parliamentarians, 3 are black, whereas 22% of the European population is black.<sup>5</sup> Similarly, in universities, women held 40.6% of academic positions in the 28 EU-countries in 2013, but few institutions have female heads (20%).<sup>6</sup> In 2016, in the USA, 27% of full-time professors are female and 4% are black.<sup>7</sup> The last example, homophily in higher education, is, in some

<sup>2</sup>Available at <https://www.catalyst.org/research/missing-pieces-report-the-2018-board-diversity-census-of-women-and-minorities-on-fortune-500-boards/>; access March 2019

<sup>3</sup>See Barnard et al. (2016)

<sup>4</sup>Available at <http://www.europarl.europa.eu/elections2014-results/en/gender-balance.html>; access March 2019

<sup>5</sup>Available at <https://www.theguardian.com/world/2018/aug/29/eu-is-too-white-brex-it-likely-to-make-it-worse>; access March 2019

<sup>6</sup><https://www.catalyst.org/research/>; access March 2019

<sup>7</sup><https://nces.ed.gov/fastfacts/display.asp?id=61>; access March 2019

ways, the most troubling. The first order effect (absence of women and people of color) is troubling enough in its own right. But there is a strong second order effect: education is a place where students learn skills and knowledge, but also where their perceptions about the world's structure are formed. An important role of schools, and universities, in particular, has been the formation of "citizens", so the social structures and group compositions to which students are exposed (including the presence or absence of homophily) can be instrumental in the nature of the citizens the education system produces. One specific relationship in which homophily might be troubling is in student-supervisor relations. Moreover, this relationship is acknowledged to be very important for future career trajectories, so biases there are likely to perpetuate the situation described above.

It is this issue, homophily in student-supervisor relationships, that I address in this chapter. Recent discussions of homophily in general identify two types: choice homophily, originating in deliberate choices of actors, and induced homophily, where the association of like with like in smaller groups is driven by the composition of the larger population from which those groups are drawn and network structures. Dis-entangling these two sources is important, as addressing homophily will demand different policies depending on its origin as chosen or induced, but it has been a challenge. One contribution of this chapter is to present a method for identifying the relative magnitudes of these two sources of homophily. Further, it develops a simple model that permits to identify the sub-groups of the population wherein homophily originates or is strongest. My empirical setting is the university system in South Africa.

Concerning social homophily, South Africa is an extreme case. Apartheid formalized homophily in the legal structure, ensuring that people were strictly segregated along ethnic or racial lines. This applied in all walks of life, including education, and the segregation in the education system may have had the deepest and most long-lasting effects. Between 1994, when apartheid ended, and 2000, social transformation to reduce homophily in universities was pursued by encouraging geographic mobility and promoting the enrollment of black and female students in former white universities. Between 2002 and 2004, a systemic change merged many university departments and created new institutions to foster diversity in the faculty composition (Herman, 2017). Reforms notwithstanding, the legacy of apartheid is still present. In the formerly white (and historically most prestigious) universities, in 2016 less than 20% of faculty with Ph.D.s was black, whereas blacks comprise 90% of the South African population.<sup>8</sup> Homophily remains an issue in higher education in South Africa.

---

<sup>8</sup> Annual Report of the Department of Higher Education (2016), "Statistics on post-school education and training in South Africa", see <http://www.ssaci.org.za/images/Statistics-on-Post-School-Education-and-Training-in-South-Africa-2016.pdf>; and Breetzke and Hedding (2018).



## 2.2 Homophily and social transformations

Homophily is an empirical regularity of our social (McPherson et al., 2001; Currarini et al., 2009), economic (Podolny, 1994; Jackson, 2005), and cultural (Barnett and Benefield, 2017) lives. Homophily is the agents' tendency to connect to those similar to themselves. In the literature similarity is conceived along many dimensions which can be divided into exogenous (like gender) or endogenous (like social status and norms). I focus on two exogenous characteristics; race and gender.

Homophily operates also in the labor market. In any workplace imperfect information plagues both sides of a hiring decision, and to reduce it agents rely on social networks. Recruiters have incentives to use their networks because they will perceive as reliable the advice coming from people they trust (i.e. their acquaintances) (Schmutte, 2016). Job seekers are likely to receive relevant information about job vacancies from their acquaintances. However, social networks tend to exhibit homophily along socio-economic dimensions. Thus, hiring decisions are often biased along the dimensions displaying homophily in social networks and they are likely to be resilient, reproducing the group structure of the organization itself (Barnard et al., 2016; Schmutte, 2016; Montgomery, 1991; Granovetter, 1973; Sherif, 2015).

Past segregation also affects the inertia within the workplace. Segregation adds additional constraints to agents' interactions and it reduces the number of linkages between demographic groups. As mentioned above, homophily can arise from choice or can be induced, but because the social space of each individual is endogenous, to disentangle the two empirically is not easy. On the one hand, the composition of the social space of an individual influences her/his choices for association; on the other hand, it emerges from her/his past choices for association. But further, individual choices do not stand in isolation, they are embedded in a society with certain norms, structures, and culture. In other words, individuals' choices are also influenced by choices made by other individuals across generations. Indeed, path dependence and stratification makes any observed behavior (of individuals or groups) a combination of induced and choice homophily (Sharmeen et al., 2014; Kossinets and Watts, 2009).

The formal definitions of choice and induced homophily evades this problem. Kossinets and Watts (2009) define induced homophily and consider choice homophily as a residual: the level of homophily that is unexplained by induced homophily. According to this definition, induced homophily is the probability of forming same-type of ties given availability of agents of same type in the potential pool (sometimes de-

defined as the risk set) (McPherson and Smith-Lovin, 1987; Kossinets and Watts, 2009) if ties are formed randomly.

This forms the basis of the strategy I use for disentangling choice from induced homophily. To my knowledge, only Kossinets and Watts (2009) study empirically induced and choice homophily. They study the network of e-mail messages of a US university for one academic year. Differently, their methodology focuses on estimating the impact of similarity and social proximity on the probability that two agents form a new tie. They find that similarity governs tie formation only when agents are not close in the social space. This result suggests that both induced and choice homophily play a role in tie formation. By contrast, I look at choice and induced homophily at two levels of aggregation in an attempt to disentangle the two empirically.

Besides the overall level of homophily, different parts of the population may have different frequency or strength of homophilous preferences (Pin and Rogers, 2016). Indeed, each group evolves differently and has different incentives in response to social transformations. In more general terms social transformations change the demographic composition of a system and the structure of agents' interactions. The way in which different population groups interact with each other is crucial to effective and stable social transformations (Smith et al., 2014). When a new generation of individuals enters a system, for example, it changes its demographic composition. Newcomers usually have fewer connections because they have been in the system for a shorter amount of time and they are on average younger. Past research finds that newcomers tend to have higher homophily levels (Mollica et al., 2003; Bramoullé et al., 2012). In addition, the workplace usually has a job hierarchy. In such a context, homophily/heterophily can be functional to the strategic behavior of the agents. In a such system, agents of the formerly dominant group are more likely to be on top of the hierarchy and each newcomer will have incentives to link with them. When this strategic behavior holds it translates into high homophily levels for newcomers of the formerly dominant group and heterophily for newcomers of the under-represented group (Mollica et al., 2003; Main, 2014; McPherson et al., 2001; Ibarra, 1992, 1997).

In the psychological and sociological literature, conflict and contact theory offer a framework which links demographic changes and homophilous behavior of the groups. In particular, a change in the demographic composition of a group changes the probability of cross-type interaction and thus preferences for association (Alderfer and Smith, 1982; Zebrowitz et al., 2008; Hewstone and Swart, 2011; Zhou et al., 2018; Barnard et al., 2016; Jost et al., 2004).

For contact theory, when the quantity and quality of interactions with members of the other group increase, homophilous preferences for same-type ties decrease. In

other words, a lack of contact with members of another group is the primary cause of homophilous decisions based on group identity (Allport et al., 1954; Pettigrew and Tropp, 2006; Hewstone and Swart, 2011; McKeown and Dixon, 2017; Stainback, 2008; Zhou et al., 2018). When the number of relevant cross-group contacts increases group identity and group homophily decreases and people will base their choices on individual characteristics rather than on group identity (Hewstone and Swart, 2011).

By contrast, for conflict theory the tension among groups of different sizes and “status” shapes group identity. This because the “in charge” group wants to maintain relevant positions or access to resources (Sherif et al., 1961; Sherif, 1966; Levine and Campbell, 1972; Brief et al., 2005). A demographic change will increase the tension among groups resulting in a hardening of attitudes and stronger homophilistic choices based on group identity.

Both theories examine important aspects of group behavior underlining the link between homophily and social transformations and its role in a system at the aggregate level.

### 2.2.1 Homophily in Higher Education and Research

The evidence of homophily in Higher Education is observable (mainly) in the demographic composition across the academic ranks and fields. At high ranks, white males still dominate academic positions.<sup>9</sup> For example, in 2016, according to the US Department of Education, only 27% of full-time professors are female, and 4% are black.<sup>10</sup> Across fields, a certain degree of segregation is evident where female and black faculty concentrate in specific disciplines (e.g. art, history, biology, medicine) and are under-represented in others (e.g. engineering, physics, computer science, and economics) (Moore et al., 2018; Weeden et al., 2017; Mullen and Baker, 2008).

A key question, to understand these imbalances, remains unresolved: is there a bias in selection committees, or is the population of applicants not representative of the population as a whole?<sup>11</sup> A way to answer this question is to understand the relative magnitude of choice and induced homophily in producing these imbalances.

The empirical literature often concentrates on choice homophily. In particular, among the vast literature documenting the lower success rates, and biases against

---

<sup>9</sup>Here academic rank refers to either as institutional prestige or as rank in the career.

<sup>10</sup>Source NCES <https://nces.ed.gov/fastfacts/display.asp?id=61>; access March 2019

<sup>11</sup>UC Berkeley was accused of (anti-female) gender bias in undergraduate admissions in the 1970s, but on closer examination, it was uncovered that most of the discrepancy in admissions rates for male and female applicants was induced by differences in the departments to which the groups applied. In fact there was evidence that choice homophily operated in favor of admitting women (Bickel et al., 1975).

women in various types of academic evaluation, several papers look explicitly at whether this bias may be driven by choice homophily among the evaluators. Results suggest that these processes are quite complex and context-specific. Bagues and Esteve-Volart (2010) find heterophilous effects in national hiring committees in the Spanish judiciary, implying that female candidates are less likely to succeed when the evaluation committee contains several women. They suggest this is because women on the committee over-estimate the quality of male applicants. On the other hand, what they observe is compatible with the opposite hypothesis, a homophilous effect. That is, the presence of women in committees may affect the voting behavior of the male colleagues such that male members increasingly favor male candidates. The latter is what Bagues et al. (2017) find in an extensive study of promotions to associate and full professor in Spain. They find that female committee members have gender-blind preferences, but when there is a woman on a committee, the men on that committee become less favorable towards female candidates. A second paper confirms a similar mechanism, with a slightly different data-set: Bagues et al. (2014) find that the success rate of a female candidate decreases by 2 percentage points for each additional female in the committee.

Additional evidence of choice homophily is in terms of research funding. Male applicants are more successful in grant applications (Van Der Lee and Ellemers, 2015), and since often adjudication committees have a male majority, this may indicate homophilous preferences there also. One recent explanation for the observation that male scientists are more successful in grant applications than are females has to do with language use (Kolev et al., 2019). Controlling for many “obvious” correlates, the authors find that a gender-gap remains in success in grant proposals submitted to the Gates Foundation (2008-17). However, when in addition, they controlled for the type of vocabulary used in the proposal, the gap was no longer statistically significant. Specifically, males tend to use more vague terms, while female use a more field-specific language.

Beyond committee decisions, scientific productivity can also be affected by homophily. Indeed, publications are often necessary for promotion and precondition for research funds. In this respect, Helmer et al. (2017) find homophily to be relevant in the peer-review process, finding that editors (male and female) have same-gender preferences.

There are no empirical studies which directly access the impact of induced homophily on the individuals. However, while biases (choice homophily) in evaluations may exist, the observed outcomes may also be induced. De Paola et al. (2015) find controlling for various individual characteristics including “quality”, women are less

likely to enter the Italian Scientific Qualification competition in fields where perceived discrimination against women is high. They also find that women are less likely to apply for promotion in fields in which females were under-represented, in other words, in fields with high levels of induced homophily.

The processes that form any observed homophily in higher education and research are numerous and complex. Further, they seem to be widespread within the different processes in which an academic is involved, such as tenure, promotion, publishing, and grant applications. One specific process that may matter a lot, but is as yet very little studied, is the selection of supervision. The student-supervisor relation is relevant in terms of education but can also have enormous effects on a student's future career.

In more general terms, the education literature finds that, at universities, an effective student-supervisor relation includes: (1) a focus on achievements and knowledge acquisition; (2) support, direct assistance, and role modeling; (3) mutual benefit; (4) a personal nature beyond strictly work-related issues; and (5) a professor having experience, influence, and achievement (Jacobi, 1991; Gersick et al., 2000; Girves et al., 2005; Crisp and Cruz, 2009). These components are particularly important (especially in early-stages) for under-represented and minority groups (Girves et al., 2005; Terrell and Hassell, 1994) and when these attributes are present in a positive sense can reduce their attrition rate (Terrell and Hassell, 1994). Role models are crucial, and mentors of the same race/gender of the students are often more effective (Bettinger and Long, 2005; Lockwood, 2006; Breda et al., 2018; Gaule and Piacentini, 2018; Rossello et al., 2020).

Thus, the student-supervisor relation, mainly for its relevance for under-represented groups, is a potential channel of social transformations if, for example, members of under-represented groups succeed in finding “good” supervisors who can further inspire their careers. At the same time, though, the formation of the student-supervisor tie is yet another place where homophily may be strongly present. I turn now to the analysis of homophily in student-supervisor relations in higher education in South Africa.

## 2.3 Data

The data originate with the South African National Research Foundation (NRF).<sup>12</sup> The NRF has a “rating system” in which scholars apply to have their research eval-

<sup>12</sup>NRF ([www.nrf.ac.za](http://www.nrf.ac.za)) is a state agency whose mission is to promote research and to develop national research capacity.

uated. In the application, each researcher has to submit her/his complete curriculum vitae, details on publications, work history, student supervision, plus individual characteristics such as race, age and gender. Almost all researchers in the country with a research oriented career are part of the NRF this because there are strong individual and institutional incentives.<sup>13</sup>

Overall, the South African university system is relatively small, having only 4034 tenured associate and full professors in 2012.<sup>14</sup>

The data include 78081 student-supervisor relationships (with 7432 total supervisors). They represent thesis supervision of bachelor (“Honours” in South Africa) (19%), master (56%) and Ph.D. (25%) students where Science Technology Engineering and Mathematics (STEM) represents 73% of the total supervisions and Social Sciences and Humanities (SSH) the remaining 27%.<sup>15</sup> The data contain information of students and professors gender, race, student level of education, year, university (39 institutions), and broad scientific field (18 categories).

The analysis aggregate data in five time-periods: 1973-1995 (3%), 1996-2000 (9%), 2001-2005 (22%), 2006-2010 (39%) and 2011-2014 (27%).<sup>16</sup> This chapter studies student-supervisor ties, looking at race and gender. Supervision relations between agents of same-types are: white-white (WW) and black-black (BB) for race, and male-male (MM) and female-female (FF) for gender. Ties between agents of cross-types are: black-white when a black student has a white supervisor (BW) or

<sup>13</sup>The NRF covers overall 30% of the South African Academics, those globally account for more than 90% of the peer-reviewed research output. The STEM fields have been part of the NRF rating system longer than have SSH fields, and in these fields, coverage is almost complete and in line with the statistics provided by the South African Council of Higher Education (CHE <https://www.che.ac.za>). Those disparities have to do to the fact that a large part of the South African higher education in SSH is concentrated in education only. Unfortunately, there are no national statistics on the composition of this part of the system, and the data are likely to do not represent that part of the system. Nevertheless, I should remark that transformations are more advanced in those areas where female and people of color were historically more involved.

<sup>14</sup>Data provided by the Department of Higher Education (DHE). Available at <https://africacheck.org/reports/how-many-professors-are-there-in-sa/>; access 2019

<sup>15</sup>Given how data are collated, if there is an incompleteness in the records, it will be at the lower levels. This is not particularly troublesome as Honours supervision involves the least student-supervisor contact. Comparing the sample composition with National statistics I can conclude that the sample is representative of postgraduate students in the country. National statistics are available at [https://www.che.ac.za/sites/default/files/publications/BS%206%20National%20Plan%20for%20Higher%20Education%20282001%29%20targets%20-%20Final\\_0.pdf](https://www.che.ac.za/sites/default/files/publications/BS%206%20National%20Plan%20for%20Higher%20Education%20282001%29%20targets%20-%20Final_0.pdf); [https://www.che.ac.za/focus\\_areas/higher\\_education\\_data/2013/participation](https://www.che.ac.za/focus_areas/higher_education_data/2013/participation); <https://www.idea-phd.net/images/Doctoral-Education-in-South-Africa-WEB-3.pdf>.

<sup>16</sup>The aggregation in 5 time-periods is done for two reasons. First, to reduce possible issues of sample biases, since NRF rating applications are sometimes not submitted yearly by individuals. The latter relates to historical reasons. The first period 1973-1995 represents the apartheid era. In that period the system was (almost) completely segregated. The second 1996-2000 is after the end of apartheid but before the reform of the university system. In that period, there was the political commitment to reduce the apartheid heritage. In the years 2001-2015 was announced and implemented a major reform of the university system. The reform created new institutions and merged some old institutions, so I isolate that period. The last period 2011-2014 is the post-reform period.

white-black for the reverse (WB), and similarly for gender: female-male (FM) or male-female (MF).

## 2.4 Preliminaries

In what follows, I will make repeated use of two things: a null model of tie-formation; and a particular network statistic. So before I start the analysis, I detail them here.

The population is made of students indicated with  $S$  and supervisors (teachers) called  $T$ . When the  $S$  and  $T$  are used with subscripts, indicating the dichotomous types  $b$  and  $w$  (or  $f$  and  $m$  for gender), they represent the relative proportions of types. For example,  $S_w$  indicates the proportion of students who are white, and so on for  $S_b$ ,  $T_w$ , and  $T_b$ .

### 2.4.1 Null Model

The null model describes the random assignment of student-supervisor ties in which a student and a supervisor are drawn randomly and independently from their respective populations. The probabilities of observing the four tie-types are:

$$\begin{aligned}
 Pr(ww) &= S_w T_w \\
 Pr(wb) &= S_w T_b \\
 Pr(bw) &= S_b T_w \\
 Pr(bb) &= S_b T_b.
 \end{aligned} \tag{2.1}$$

For example,  $Pr(ww)$  is equal to the proportion of white students multiplied by the proportion of white supervisors ( $S_w T_w$ ). The four probabilities represent the random mixing which is the probability of observing a certain tie-type given population availability and type-blind tie formation.

#### Null model with permutations

To account for the small sample bias and other confounding embedded in the network, I create a “corrected” null model using permutations. The null model with permutations is the result of tie-types given by repeated permutations of the existing ties. So the probability of observing any of the four tie-types is given by permutations results, as follows

$$\begin{aligned}
Pr(ww) &= \frac{1}{J} \sum_{j=1}^J W_{perm_j} W \\
Pr(wb) &= \frac{1}{J} \sum_{j=1}^J W_{perm_j} B \\
Pr(bw) &= \frac{1}{J} \sum_{j=1}^J B_{perm_j} W \\
Pr(bb) &= \frac{1}{J} \sum_{j=1}^J B_{perm_j} B.
\end{aligned} \tag{2.2}$$

Where  $J$  is the number of permutations,  $j$  is a realization of the permutation, and  $perm_j$  indicates a random assignment of students to supervisors. The permutation preserves the degree sequence of supervision. In other words, it retains the number of supervisions that each professor performs at the individual level and randomly assigns the students.

### 2.4.2 Homophily and Assortativity

To measure homophily I use the assortativity coefficient (Newman, 2003). It is a standard measure used to characterize this aspect of social networks (Johny et al., 2017). The assortativity coefficient is in general terms

$$r = \frac{\sum_i e_{ii} - \sum_i a_i b_i}{1 - \sum_i a_i b_i}. \tag{2.3}$$

$i$  indicates the types ( $w$  and  $b$ );  $e_{ii}$  are the proportions of same-type of ties (WW and BB in this case) and  $a_i$  and  $b_i$  are the proportions of each type in the population  $a$  and  $b$  ( $S$  and  $T$  in this case). Here:

$$Ass. = \frac{(BB + WW) - (S_b T_b + S_w T_w)}{1 - (S_b T_b + S_w T_w)}. \tag{2.4}$$

Where BB and WW are same-type ties proportions,  $S_b$  is the fraction of black students,  $T_b$  is the fraction of black professors, and  $S_w$  and  $T_w$  are, respectively, those of white students and professors. Intuitively, the assortativity coefficient measures the gap between observed same-types ties and those predicted by the null model (eq.2.1). To reduce the issue of small sample bias and to control for various confounding factors embedded in the network structure I use in section 2.8 a modified version of the assortativity coefficient. I replace the expected value with the permuted values as described above. This gives a frequency distribution for the assortativity measure rather



than a point observation.

With those preliminaries in place, I turn now to the analysis.

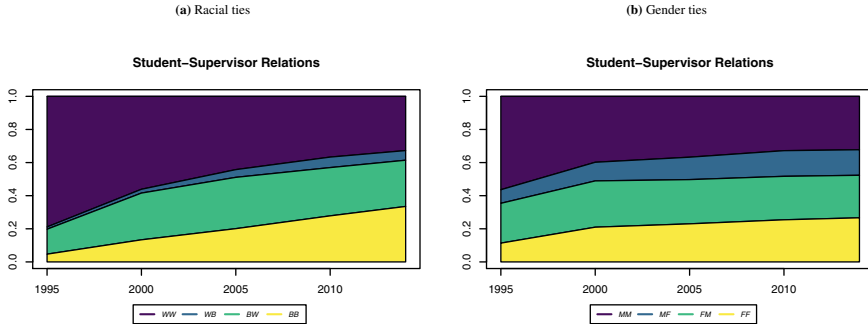
## 2.5 Student-Supervisor ties

**Table 2.1:** Student supervision composition. Where  $w$  indicates white,  $b$  black,  $m$  male,  $f$  female,  $S$  represents students, and  $T$  supervisors. *Ass.* is an abbreviation for assortativity. Expected values of random mixing (eq. 2.1) are in parenthesis. Columns 1 and 2 should be treated carefully as both have few observations. Column 1 represents the apartheid era and column 2 is before the reform of the university system.

	1973-1995	1996-2000	2001-2005	2006-2010	2011-2014
Populations					
$S_w$	0.80	0.58	0.49	0.43	0.39
$S_b$	0.20	0.42	0.51	0.57	0.61
$T_w$	0.90	0.80	0.71	0.61	0.56
$T_b$	0.10	0.20	0.29	0.39	0.44
Supervisions					
$WW$	0.79 (0.72)	0.56 (0.46)	0.44 (0.35)	0.37 (0.26)	0.33 (0.22)
$WB$	0.01 (0.08)	0.02 (0.12)	0.05 (0.14)	0.06 (0.17)	0.06 (0.17)
$BW$	0.15 (0.18)	0.28 (0.34)	0.31 (0.36)	0.29 (0.35)	0.28 (0.34)
$BB$	0.05 (0.02)	0.13 (0.08)	0.20 (0.15)	0.28 (0.22)	0.33 (0.27)
<i>Ass.b/w</i>	0.38	0.31	0.29	0.32	0.34
Populations					
$S_m$	0.65	0.51	0.50	0.48	0.48
$S_f$	0.35	0.49	0.50	0.52	0.52
$T_m$	0.74	0.67	0.62	0.57	0.56
$T_f$	0.26	0.33	0.38	0.43	0.44
Supervisions					
$MM$	0.56 (0.48)	0.40 (0.34)	0.37(0.31)	0.33 (0.27)	0.32 (0.27)
$MF$	0.08 (0.17)	0.11 (0.17)	0.14 (0.19)	0.15 (0.21)	0.15 (0.21)
$FM$	0.24 (0.26)	0.28 (0.33)	0.27 (0.31)	0.26 (0.30)	0.26 (0.29)
$FF$	0.11 (0.09)	0.21 (0.16)	0.23 (0.19)	0.25 (0.22)	0.27 (0.23)
<i>Ass.m/f</i>	0.23	0.21	0.2	0.16	0.18

The stack plot of tie-type composition over time underlines the social transformation of the university system (fig. 2.1). In particular, it shows the growth of cross-type ties and the growth of black (female) participation. In more details, descriptive statistics on populations, supervisions, and the network assortativity coefficient are in Table 2.1. The table underlines the transition of the system: from a system dominated by white males to one showing more diversity. The presence of black (female) students increases from  $S_b = 20\%$  ( $S_f = 35\%$ ) in 1973-94 to  $S_b = 61\%$  ( $S_f = 52\%$ ) in 2011-14. A similar, though less striking transformation occurs in the composition of the faculty: the percentage of black (female) professors ( $T_b$  and  $T_f$ ) increases from 10% (26%) to 44% (44%) over the same periods.

**Figure 2.1:** Stack plot of Student-Supervisor relations 1973-2014. Ties are labels ordered as student, supervisor. Racial ties *WW*, *WB*, *BW*, *BB* (2.1a) and Gender ties *MM*, *MF*, *FM*, *FF* (2.1b). For 5 time periods: 1973-1995; 1996-2000; 2001-2005; 2006-2010; and 2011-2014



The compositional changes of the students and the supervisors provoked an increase in cross-type of relationships, evident in the second parts of each panel. In the period of apartheid, 1973-1995 white professors have 94% of the total supervisions, and supervisors are 80% male. In the last period of the sample, 2011-2014, by contrast, white and male professors supervise respectively 61% and 58% of the students. Same racial supervision ties (WW and BB) represent the 84% of the total in 1973-1995, they fall to 66% in 2011-2014. Similarly, same-gender supervisions (MM and FF) fall from 67% in 1973-1995 to 59% in 2011-2014. In each cell, the parenthesis indicates the proportion of supervisions of that type expected under the null model (eq. 2.1). Thus, besides the increase of cross-type of ties over time, ties of white (male) students with black (female) supervisors are relatively rare even today.

Newman's assortativity coefficient in the final row of each panel shows positive values for race and gender assortativity. A positive coefficient indicates a tendency (at the aggregate level at least) for agents to connect with those of same-type. This provides prima facie evidence of homophily.

Race assortativity decreases before the university reform (2001-2005) but, after, it increases again. Instead, gender assortativity appears more stable.

### 2.5.1 Future Academics

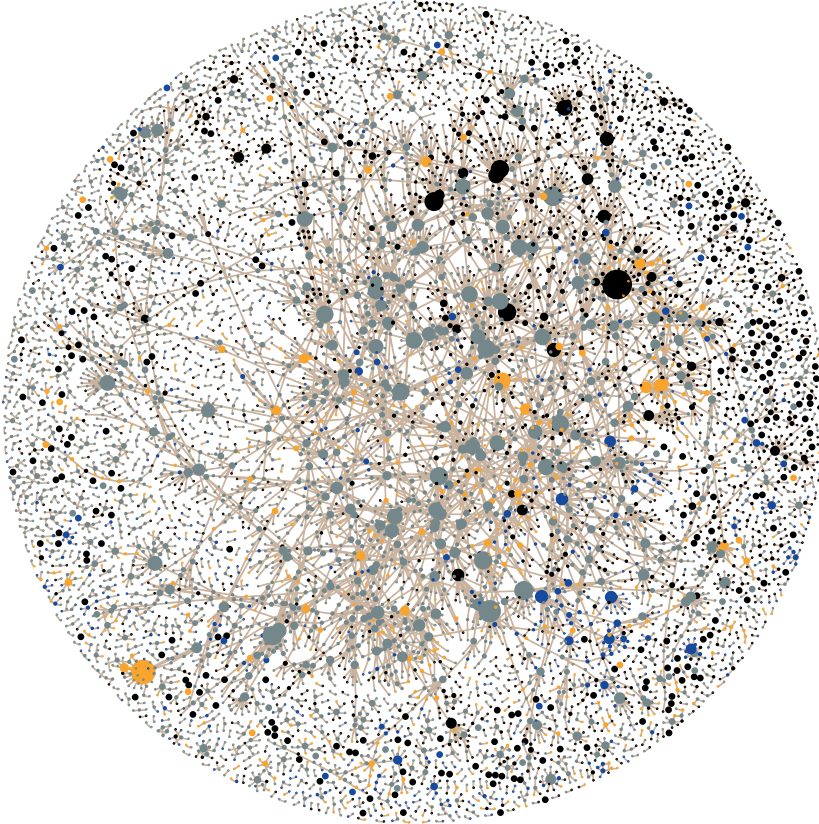
In the data I identify students who continue their career as (research-active) scholars in a South African university. The NRF database is an extensive database with many tables containing different information of all individuals who applied to NRF between 1970 to 2014. In this chapter, I identify future academics with the following procedure. I use the data of student-advisor records described before and I matched them with university employment records. First, I started from the database contain-

**Table 2.2:** Student supervision composition subsample of Future Academics. Where w for white, b for black, m for male, f for female, S for students, and T for professors. Ass. is abbreviation for assortativity. Expected values of random mixing are in parenthesis. To note: column 1 and 2 should be treated carefully, both have few observations. Column 1 represents the apartheid era and column 2 is before the reform of the university system.

	1973-1995	1996-2000	2001-2005	2006-2010	2011-2014
Populations					
$S_w$	0.79	0.61	0.50	0.46	0.40
$S_b$	0.21	0.39	0.50	0.54	0.60
$T_w$	0.93	0.87	0.75	0.64	0.58
$T_b$	0.07	0.13	0.25	0.36	0.42
Supervisions					
$WW$	0.78 (0.73)	0.60 (0.53)	0.46 (0.38)	0.40 (0.29)	0.35 (0.23)
$WB$	0.01 (0.06)	0.02 (0.08)	0.03 (0.13)	0.06 (0.17)	0.05 (0.17)
$BW$	0.17 (0.20)	0.29 (0.34)	0.33 (0.38)	0.26 (0.35)	0.24 (0.35)
$BB$	0.04 (0.01)	0.09 (0.05)	0.17 (0.13)	0.27 (0.19)	0.36 (0.25)
$Ass.w/b$	0.28	0.26	0.26	0.35	0.44
Populations					
$S_m$	0.67	0.51	0.51	0.46	0.40
$S_f$	0.33	0.49	0.49	0.54	0.60
$T_m$	0.79	0.73	0.67	0.61	0.58
$T_f$	0.21	0.27	0.33	0.39	0.42
Supervisions					
$MM$	0.60 (0.53)	0.40 (0.37)	0.37 (0.34)	0.33 (0.28)	0.29 (0.23)
$MF$	0.07 (0.14)	0.11 (0.14)	0.14 (0.17)	0.12 (0.18)	0.11 (0.17)
$FM$	0.23 (0.26)	0.32 (0.36)	0.30 (0.33)	0.30 (0.33)	0.29 (0.35)
$FF$	0.10 (0.07)	0.17 (0.13)	0.19 (0.16)	0.24 (0.21)	0.30 (0.25)
$Ass.m/f$	0.25	0.13	0.11	0.15	0.21

ing employment records. Here for each individual is presented the career progression, including graduation year, broad scientific field, and university of employment. First, I disambiguated all names and surnames, and I attached a unique user id to all uniquely identified agents. Then I matched Students records with records on the employment database matching records by name, surname, university, scientific field and degree year. I manually checked a random sample of 100 individuals to access the quality of the matching searching on-line information of these individuals. The students identified in this way are the future academics that after graduation applies to NRF between 1970 to 2014. This procedure has one limitation: the observation period after graduation is not fixed for all agents. However, the students in the last period 2011-2014 are about 27% of the total and NRF gives strong incentives also for young scholars to apply to be rated. Future academics are an interesting sub-group, as their super-

**Figure 2.2:** Student-Supervisor network for Future Academics 1973-2014. The size of the nodes are proportional to node degree. The colour of the nodes represents ethnic groups black are black, blue are coloured, orange are Indians and grey are white



visors may have identified them as promising students. They constitute 13% of the sample of students. I consider the future academics sample important to understand the entry process of researchers into the system. Table 2.2 shows summary statistics for this sub-population of students. The population composition and supervisions of the sub-sample are in line with the overall population, though social transformation appears slightly lower along racial lines. In figure 2.2 I show for the sample of future academics the network of student-supervisor relations where the colours of the nodes represent ethnic groups and the size of nodes their number of connection (node degree). The figure shows to some extent the degrees of segregation in the system. Indeed, in the giant connected component in the middle the more central agents with

a higher number of connections are mostly whites (grey nodes). All around, in a more peripheral position are visible clusters of black (black, blue, orange) students and advisors.

## 2.6 Methodology

A system has various layers of organizations or (sub-)units; for this, the key to isolating choice homophily in a system is to use different levels of aggregation. The university system is a collection of the faculties, each one organized into departments. In principle universities and departments interact with each other, but supervision happens (mostly) within a department. Indeed, institutional constraints govern student-supervisor tie formation, and each university department has the duty and the incentive to provide internal supervision for their students. In other words, the department is the risk set, the meeting pool of the agents likely to form ties together. University departments represent the constraints to agents interaction which create induced homophily within the system.<sup>17</sup> These organizational “limits” to interaction express norms, specialization and the division of labor. For example, lab equipment available in few universities could limit supervision and affects the formation of student-advisor relation. This chapter includes that in the broad definition of institutional constraints. Indeed, equipment, funding, and facilities are part of the institutional arrangement of a university system. This type of constraints produces induced homophily in the university system. I control for induced homophily isolating choice homophily using the department level (randomization) permutation. The randomization at university department level controls for all type of institutional constraints (e.g. department demographic composition, professor popularity, funding, norms, lab equipment, facilities).<sup>18</sup>

Thus, to remove most of the induced homophily, I restrict attention to the students and supervisors within a single department — any student within a department can

<sup>17</sup>In some cases, there are demonstration effects at play. This happens when students choose a certain department because there is a specific demographic composition. In this situation, I argue that what it may appear as “choice homophily” of students is instead induced by the system. Let’s consider this thought experiment. In a hypothetical system where all groups are fairly represented in all (sub-)units or ranks, that will not happen. A student will enter the system and chose her/his major and only then, after knowing professors in her/his department will make a choice for association. In that case, after the opportunity provided by the system to be exposed to diversity, I can really consider her/his choice as “free” and not induced. In this line of reasoning, I consider this sorting effect into disciplines or department as induced by the system.

<sup>18</sup>A caution remark should be considered here. Within one university department, not all have equal access to the scientific equipment in the laboratory. A star scholar might access equipment more than his co-workers. Related to this issue, I made an implicit assumption in my methodology that is: a star scholar attracts more students then a colleague (i.e. she/he displays preferential attachment in supervision). However, this might be a limitation that could be overcome with the use of collaboration data. Part of it is accounted in Chapter 3 advisor past average productivity is added as a control. An additional limitation is in the way I am forced to define university departments.

take any faculty within that department as a supervisor.<sup>19</sup>

The first part of the chapter empirically investigates homophily in its two components choice and induced. I use a permutation technique repeated at the system and the department level of aggregation. I use the pair university name and broad scientific field, to identify university departments.<sup>20</sup> The permutation results give the null model of random type-blind ties. Then, using permutation results, I compute homophily levels in the two cases. The measure of homophily is a modified version of the assortativity coefficient (see section 2.4) and examines at the gap between the null model with permutations and observed ties (eq. 2.5 below).

Homophily measured at the system level comprises both choice and induced. The one measured at the department level eliminates (much of) the homophily induced by the system and represents choice homophily.

The second part of the paper, section 2.9 below, I investigate where choice homophily originates. I use department-level data to measure the strength of homophilous preferences in the populations of students and professors. In particular, I develop a model to estimate homophilous preference in the four (sub-)groups: black and white students and black and white professors.

In section 2.10, I further discuss the model showing how measures of homophily can be prone to misleading interpretations in the presence of social transformations.

## 2.7 Permutation Test

I perform a permutation test creating 95% confidence intervals of the null model with permutations (see section 2.4 eq. 2.2) at the aggregate level, and then at the department level. This to test whether observed tie-types are more (less) likely than a model of random type-blind tie formation would predict. If there is no homophily in student-supervisor tie formation, the observed tie-types should lie within the confidence intervals of the permutation test.

The permutation test at the aggregate system level is constructed as follows. For each period, it permutes the students 100 times and randomly assigns students to “supervision slots”. In other words, it retains the students and professors actual population, and the total of supervisions each supervisor performs (keeping fixed the supervisor’s degree sequence) and randomizes the students.<sup>21</sup> From this set of 100 permu-

<sup>19</sup>It must be noticed that for example, a physics departments might include both theoretical and experimental physicists. An experimenter would not supervise a student doing a theoretical thesis. So there will remain some induced homophily for which I cannot correct. This is particularly true given the way I am forced to operationalize “department”, as I discuss below.

<sup>20</sup>The full list of university names and fields are in the Appendix A.2.

<sup>21</sup>The procedure preserves the degree sequence of student supervision, randomizing the demographic

tations I compute averages and 95% confidence intervals of tie-types, and I compare them with observed ones.

At the system level, part of the gap between observed tie-types and the permutation test is due to the implicit assumption of the permutation, which is that each professor can supervise anyone: a physicist in Cape Town can supervise a student of law in Limpopo. In the university system there are many constraints (e.g. geographic, disciplinary) which prevent the formation of such relation, this makes part of the observed divergence induced. This is what I refer to as the homophily induced by the system.

I remove induced homophily repeating the permutation test at the department level. I use the same procedure, but I restrict the permutation to run within the university department. Due to data limitation, my proxy of a university department is the couple university name and scientific field. In particular, there are 39 South African institutions and 18 scientific fields.<sup>22 23</sup>

I run the permutation tests at the system and the department level for the whole sample and the following sub-samples of the population: identified students who enter academia (Future Academics), STEM, SSH, top institutions<sup>24</sup>, and Ph.D. students (the last two are in Appendix A.1)

### 2.7.1 Permutation Test results

Figure 2.3 shows the results of the permutation test at the system and the department level for different sub-samples of the data. The left column of panels treats the entire country as one integrated system, imposing no constraints on who can supervise whom (system level permutations - without institutional constraints). The right column of panels constrains supervisions to take place within one “department” (department level permutations - with institutional constraints).

Results for different (sub-)samples are consistent.<sup>25</sup>

Figure 2.3 underlines that student-supervisor ties display racial (gender) homophily (results for gender are in figure 4 in Appendix E). In particular, observed proportions of tie-types (solid lines) are outside the 95% confidence interval of the permutation test (dashed lines). In other words, same(cross)-type of ties are more (less) likely than

---

composition of the ties. It controls for preferential attachment and other confounding factors, like the willingness of supervisors to accept students, budget constraints, and faculty practices and norms.

<sup>22</sup>Institutions and fields are listed in the Appendix A.2

<sup>23</sup>The proxy for university departments is probably too broad to get precisely the right level of disaggregation. This is a limitation of the data.

<sup>24</sup>Top institutions are: Cape Town, Pretoria, KwaZuluNatal, Stellenbosch, Rhodes, Witwatersrand, and Western Cape University.

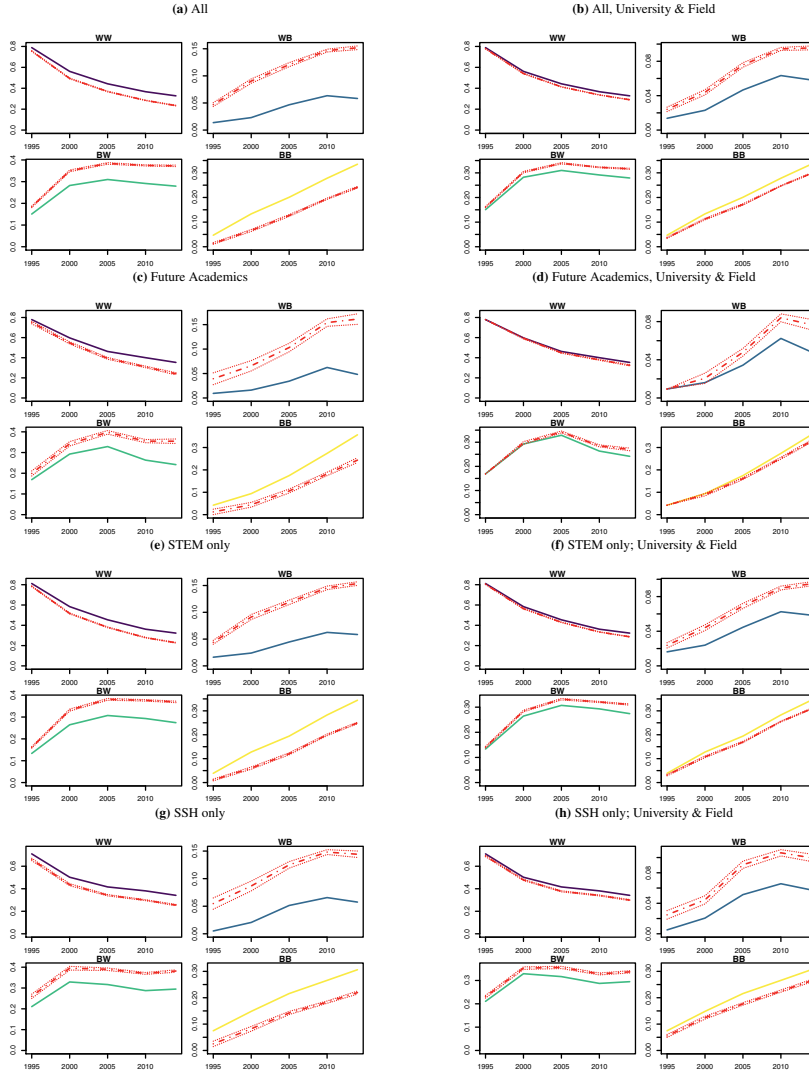
<sup>25</sup>As a robustness check, I performed the analysis the permutation in each year and then aggregating into the five periods and the results are consistent.

a process of random type-blind tie formation would imply.

The results comparison of the permutation tests at the system and department level (figures 2.3 and 4 from the right column to the left) shows that the gap between observed tie-type and the null model is lower when institutional constraints (department level permutation) are taken into account. Though it remains statistically significant, this highlights the presence of a large amount of induced homophily. In the following section, I measure more directly the relative magnitude of induced and choice homophily.



**Figure 2.3:** Permutation tests for student supervisor data 1973-2014. The permutation is done keeping fixed the number of ties per supervisor and permuting the students. Permutation without constraint (left) and with institutional constraints (right). The Permutation is repeated 100 times for each of time period. The results of the permutation are plotted using dashed lines with two standard deviations on either side of the mean shown in the same color. Solid lines in each plot show the proportions observed in the data. Tie-type labels (BB, BW, WW, WB) are read as student-supervisor.



## 2.8 Choice and Induced homophily

In this section, I use the system and department level permutation to separate choice homophily from that induced by the system.

I measure homophily modifying Newman's assortativity coefficient (in eq. 2.4) as follows. The Newman coefficient considers the gap between the proportions of same-type ties observed and expected values of random mixing in the population; in contrast, I compare observed same-type of ties with the distribution of permutation results. The measure becomes

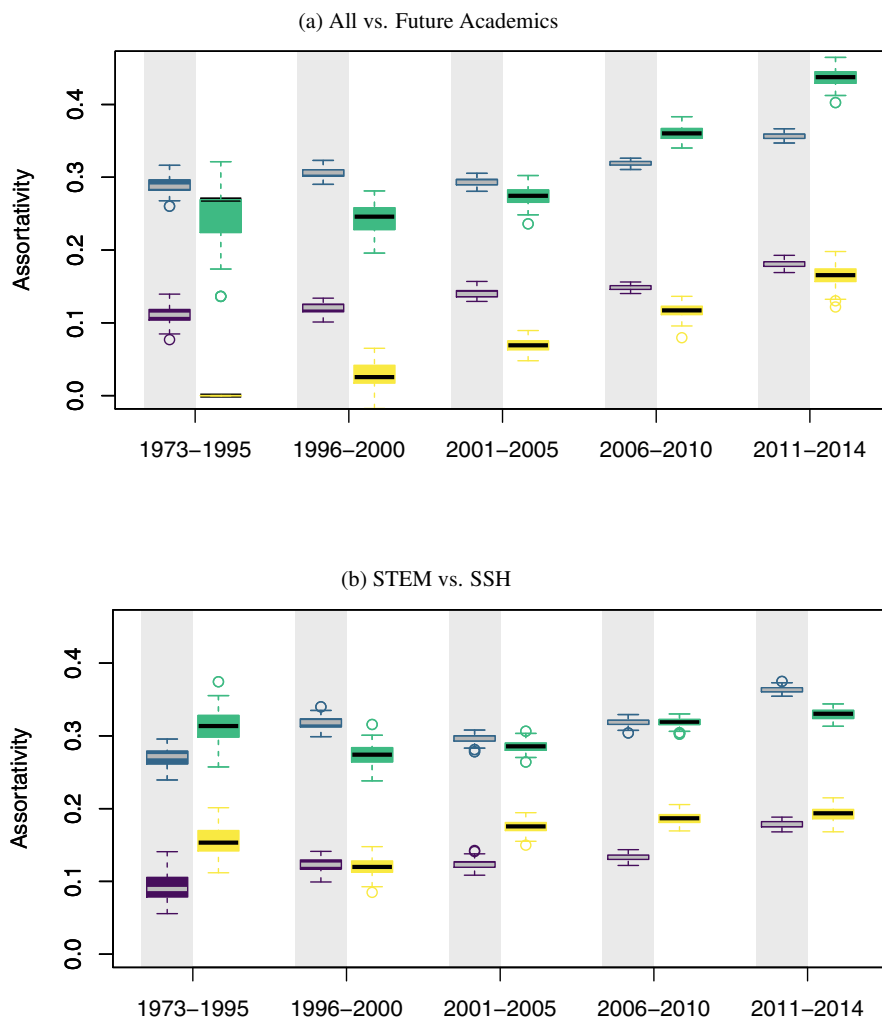
$$Ass_j = \frac{(BB_{obs} + WW_{obs}) - (B_{perm_j}B + W_{perm_j}W)}{1 - (B_{perm_j}B + W_{perm_j}W)}. \quad (2.5)$$

$j$  is a realization of the 100 permutations.<sup>26</sup>  $BB$  and  $WW$  are proportions of the same-type tie, and  $obs$  and  $perm$  mean observed and permutation results. I present homophily computed in this way using box-plots.

---

<sup>26</sup> $B_{perm_j}B$  indicates that students are randomly assigned to supervisors, maintaining the degree sequence for supervisors.

**Figure 2.4:** Induced and Choice Homophily comparison looking racial ties in student supervisor relations 1973-2014. I compare results of assortativity of different sub-samples: All vs. Future Academics (a), STEM vs. SSH (b). Each panel contains 4 series; In Fig.(a): aggregated (upper series) versus department level (lower series) crossed with total sample (white background) versus future academics (grey background). And similarly for Fig. (b). Upper series are system-level permutations, including both induced and choice homophily; lower series are department-level permutations, excluding much of the induced homophily. Each box plot represents 100 permutations and associated assortativity calculation, as described in section 2.4.2, and equation 2.4.



In each column and period figure 2.4 shows assortativity at the system level in the upper boxes and assortativity at the department level in lower boxes. In the lower boxes the institutional constraints are included, I have constrained supervisions to take place within a department, which approximates a situation of zero induced homophily—in principle any professor can supervise any student. The lower boxes thus approximate choice homophily and the difference between upper and lower boxes

is an estimate of induced homophily.

Figures 2.4(a) and 2.4(b) show choice and induced racial homophily. Figures 2.4(a) compares the whole sample (grey background) with future academics (white background). While Figures 2.4(b) compares STEM (grey background) with SSH (white background). Both figures underline that at the system level homophily (upper boxes) is roughly double that at the department level (lower boxes). Thus a large part of overall observed homophily is induced by department composition. Further, choice homophily (lower boxes) exhibits a slightly increasing trend. A similar trend could be partially an artefact of social transformations taking place in the system. I explore this further in section 2.10.

The Comparison between the whole sample and the sub-sample of future academics shows additional insights. The results underline that future academics experience larger institutional constraints and so higher induced homophily. In particular, after 2001-2005, the assortativity measured for future academics in figure 2.4(a) (white background - upper boxes) exhibits higher levels compared with the whole sample (grey background — upper boxes) while by contrast, choice homophily (lower boxes) has similar levels. Such a result could be compatible with demonstration effects along racial lines in (sub-)disciplines. The latter creates a correlation between sub-field and racial composition of the academics in it. Demonstration effects work in the following way: when students choose an academic discipline, they ask whether “people like me” prosper in that discipline. Demonstration effects and absence of role models are often an explanation for the under-representation of women in STEM subjects (Steele and Ambady, 2006; Bian et al., 2017; Blau and Kahn, 2017; Cheryan et al., 2017).

Another important indication is given comparing STEM and SSH fields. The comparison in figure 2.4(b) globally finds similar levels of homophily, but different composition. In STEM (grey background) induced homophily appears stronger. Indeed the distance between total homophily and choice homophily is larger in STEM than in SSH. If a demonstration effect is at work, this suggests it is stronger in STEM than in SSH.<sup>27</sup>

In the Appendix E figure 6 shows the analogous results for gender. In general homophily along gender lines tends to be lower than homophily along racial lines. However, estimates and time-trends of choice homophily are very similar for race and gender.<sup>28</sup>

<sup>27</sup>This is consistent with the large empirical evidence in the education literature underlining the importance of same race (gender) role models. Additionally, there are also evidences that parents educational attainments largely correlates with the educational outcomes of children. Those children may have lower needs of role models because they have it within their family.

<sup>28</sup>The low numbers of black females in the sample make unfeasible to look at the intersection between race and gender. Black scholars and students with a research-active career are mostly male. Indeed, statistics from the South African DHE on professorship and students enrolments underline that most of the racial

Globally, the results show that the composition of the departments induces race-based and gender-based homophily largely. In other words, university and field constraints to interaction generate half of the homophily observed in the system. Further, homophily, have a different composition in STEM and SSH. This suggests an unequal distribution of social transformation in the South African university system. This is consistent with the observation in Herman (2017) who suggest a knowledge divide created by apartheid and still present. The university reform of 2002-2004 in South Africa partially accounts for that. The reform reduced the number of universities from 36 to 23 and operated with mergers to redistribute resources and to remove the racial inequalities inherited from apartheid (Herman, 2017).

## 2.9 Where does choice homophily originate?

In the previous sections, I measured choice homophily in the aggregate. This assumes that in the population all have the same degree of homophilous preferences. Given the history of the country, this might not be the case. The new methodology in this section presents a way for distinguishing homophily levels among the different groups. First, I create a model of type-blind tie formation which inserts for each sub-population of students and supervisors a specific level of homophilous preferences. Second, I estimate for each sub-population the proportion of agents with homophilous preferences at the department level.<sup>29</sup> Lastly, I discuss the need for caution interpretation of the assortativity coefficient in a transition phase of social transformations.<sup>30</sup>

Intuitively, the model includes in each sub-population a certain fraction of agents with strictly homophilous preferences (that form only same-type ties) and the remaining is completely “color-blind” (form ties at random).<sup>31</sup> Four are the sub-population groups: white and black (male and female) professors and black and white (male and female) students. I apply the model at the department level, thus eliminating induced homophily.

---

inclusion happens with the introduction of black males. Although, the intersection between race and gender is crucial. I address it in a smaller sample which concentrates on Ph.D. STEM graduates where the sample is more complete (Rossello et al., 2020). Not surprisingly, I find a high correlation between race homophily and gender homophily. Interestingly, the results show that cross-racial couples attenuate the gender productivity gap.

<sup>29</sup>In Appendix A.3, I use a simple econometric model as a robustness check.

<sup>30</sup>Homophily values are likely to be heterogeneous across universities and departments. In principle, that is present in the model I present. However the data are not strong enough to include that aspect in the estimates, so I assume that university-level effects are not present.

<sup>31</sup>The model could as well be explicated in terms of “tendencies” for (non-)homophilous tie formation at the individual level. The analysis would be identical.

### 2.9.1 A Model of homophilous preferences of the different groups

In this section, I describe the model of random tie formation of student-supervision. I use it to estimate the homophilous preferences in the four population groups. In particular, Both students and professors are of two types  $w$  or  $b$ .

Students and professors form supervision ties according to a two-stage process that I model as a probability tree. First, I draw a student and then the professor. Students-supervisor ties are formed by drawing one member randomly with replacement from each population.

In many disciplines (particularly at the bachelor and master level) is the student who approaches the professor asking for supervision.<sup>32</sup> This mechanism is mimicked by the proposed pairing of students and supervisors, where the student is drowned first.

The possible outcomes are the probabilities of observing same- and cross-types of tie. When all agents are type-blind, according to equations 2.1, only relative sizes of groups matter, in this model, though, I add homophilous preferences.

The model has a simplifying assumption. Each sub-population has a specific level of homophilous preferences. Formally this level represents the frequency of agents in each group with strictly homophilous preferences. I assume that strictly homophilous individuals form ties only with those of same-type and refuse cross-type of ties<sup>33</sup>

In each sub-populations, there are agents with homophilous preferences who form only same-type of ties and type-blind agents who link at random in their potential pool. According to the model, the probability of  $WW$  ties is the sum of the probability that a homophilous  $w$  student ( $S_w h_{sw}$ ) is extracted plus the probability that a type-blind  $w$  student links at random with a  $w$  professor ( $(S_w(1 - h_{sw}) \frac{T_w}{1 - h_{tb} T_b})$ ). The other probabilities of tie-types are

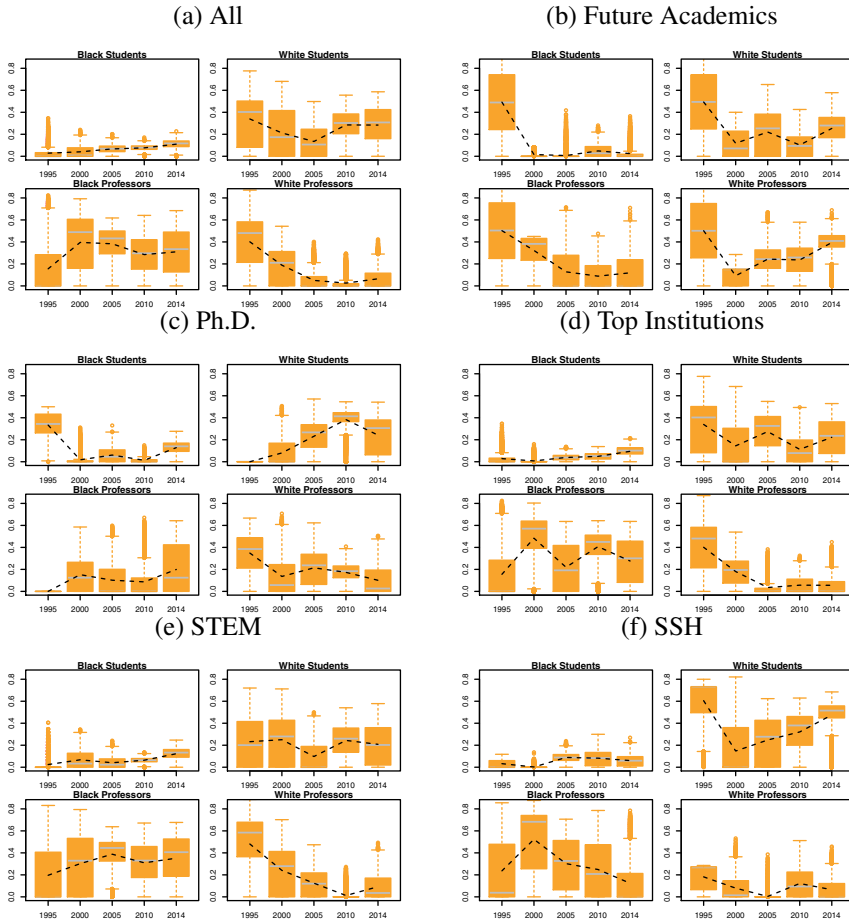
$$\begin{aligned}
 Pr(ww) &= S_w \left( \frac{T_w + h_{sw}(1 - h_{tb})T_b}{1 - h_{tb}T_b} \right) \\
 Pr(wb) &= S_w T_b \left( \frac{(1 - h_{sw})(1 - h_{tb})}{1 - h_{tb}T_b} \right) \\
 Pr(bw) &= S_b T_w \left( \frac{(1 - h_{sb})(1 - h_{tw})}{1 - h_{tw}T_w} \right) \\
 Pr(bb) &= S_b \left( \frac{T_b + h_{sb}(1 - h_{tw})T_w}{1 - h_{tw}T_w} \right).
 \end{aligned} \tag{2.6}$$

$h_{sw}$  and  $h_{sb}$  are respectively proportions of agents with homophilous preferences in

<sup>32</sup>In some disciplines for Ph.D.s in particular professors do approach students. In the analysis, Only 27 % of the data are Ph.D. supervisions.

<sup>33</sup>Skvoretz (2013) highlights the importance of including attraction to similar and repulsion from dissimilar in a model of group behavior in terms of homophily. The model includes both mechanisms.

**Figure 2.5:** Race Homophily of the different groups 1973-2014. The estimation is done minimizing the average relative entropy of tie-types between observed and predicted from equations 2.6. I use data at the department level with 10000 samples with replacement of department observations. Black dashed lines are averages. Faculties with fewer than 10 observations and without population variability are excluded.



the population of white and black students.  $h_{tW}$  and  $h_{tB}$  are those of white and black supervisors.

I use equation 2.6 to estimate the different  $h$ s of 4 sub-populations at the department level each period.

## 2.9.2 Choice Homophily of different groups

In this section, I estimate the  $h$ s of the model in the previous section. For a vector of  $h = (h_{SW}, h_{tW}, h_{SB}, h_{tB})$  the model 2.6 predicts the composition of type-ties for a given department population composition.

For the estimation I use department level data for each time period of sub-population

proportions and tie-types.<sup>34</sup> Fully black or white departments and those with less than 10 supervisions per period are removed. In particular, I include departments which satisfy  $0.1 < T_w, T_b, S_w, S_b < 0.9$  and have more than 10 supervision in a time period.

I use the predictions of equation 2.6 to estimate the four  $h$ s

I apply a bootstrap technique for the estimation (Efron and Tibshirani, 1986). I generate 10000 bootstrap samples  $U' = \{u'_1, \dots, u'_{10000}\}$ , for each period. Every  $u'_k \in U'$  has size  $M(t)$  and it is drawn from the set of university departments.<sup>35</sup> I estimate  $h = (h_{sw}, h_{tw}, h_{sb}, h_{tb})$ , for each bootstrap sample  $u'_k$ , by minimizing the average relative entropy. In particular I follow the formula (Kullback and Leibler, 1951)

$$\overline{D_{KL}(p||q)} = \frac{1}{J} \sum_{j \in J} \sum_{i \in I} p(i|j) \log \frac{p(i|j)}{q(i|j)}. \quad (2.7)$$

Where  $I = \{ww, wb, bw, bb\}$ ,  $J = \{\text{departments in the bootstrap sample } u'_k\}$ ,  $p$  are the empirical probabilities and  $q$  are the model predictions.

The procedure minimizes the average loss of information between observed proportions of tie-types and those predicted by the model in equation 2.6 at the department level. Each time it obtains an estimate of the fraction of agents with strictly homophilous preferences in the 4 sub-populations. I report the bootstrap distributions for the estimates.<sup>36</sup>

Figure 2.5 shows the estimation of the racial homophilous preferences (see figure 8 in Appendix E for gender). The results highlight that white (male) students have the highest estimated homophilous preferences. Among professors are black (female) which display the higher levels. In contrast, in the sample of future academic students, black professors have lower homophilous preferences than white professors, especially in the most recent period, 2011-2014.

It must be noticed that while black (female) professors display homophilous preferences, black (female) students do not. Thus, since it is often the student who asks a professor to supervise her/his work, black (female) professors may face additional induced homophily that I do not capture. Under this hypothesis, the students drive tie formation, and by this mechanism, they may transmit homophily.<sup>37</sup> Moreover, when I

<sup>34</sup>I account for various confounding embedded in the network (like preferential attachment, supervision capacity, etc.) repeating the estimation using the proportion of each sub-population of students and supervisors computed from supervision ties instead from actual individuals. This to include the fact that some professors may attract more students than others. The results are consistent.

<sup>35</sup> $M(t)$  is the number of departments observed in the data in period  $t$ .

<sup>36</sup>For the minimization I used the limited memory algorithm for bound constrained optimization as in Byrd et al. (1995), under the optim R function. This because the  $h$ s are probabilities thus are bounded between 0 and 1.

<sup>37</sup>This is likely to be the case in bachelor and master students, but probably less so in PhDs. I check it repeating the analysis in the sub-samples of only PhD students. Those results show little differences compared to estimates in the whole sample. Suggesting that this hypothesis is probably negligible in the



concentrate on future academics, limits to supervision experienced by black (female) professors could be stronger since, there, white professors have high homophilous preferences.

## 2.10 A simple model with changing populations and unchanging homophily

Results in section 2.8 show that choice homophily is increasing over time. Given the South African context, this would be somewhat distressing if true. I point out in this section that this can be partially a composition effect of changes in the sup-population. I illustrate this effect below.

In a simple model with  $S_w = T_w$  and the same fixed level of homophilous preferences  $h_{sw} = h_{sb} = h_{tw} = h_{tb} = 0.3$  in all sub-populations. I start with a population of 500 white students and 500 white supervisors; they are all of the same type meaning that initially  $S_w = T_w = 1$ . Progressively I introduce in such population the other demographic type.  $S_b$  and  $T_b$  which was zero at the beginning starts to rise.<sup>38</sup> I present the results as a function of the number of blacks, ( $N_b$ ), in the system.<sup>39</sup> I make supervision assignments using the model in section 2.9.1 and then calculate proportions of tie-types and assortativity mixing.

I show the behavior of the model under the described conditions of unchanged tastes-for-similarity over time in figure 2.6. Figure 2.6(a) shows on the x-axis the number of blacks in the population and on the y-axis the proportion of tie-types. It relates to figure 2.3. The figure underlines that in the system, as the number of blacks ( $N_b$ ) increases the distances between “observed” tie-types and the null model of type-blind formation (eq. 2.1) initially rises.

Figure 2.6(b) relates to figure 2.4 and shows the Newman assortativity coefficient as a function of the number of blacks in the population. It shows that assortativity makes a steep rise, before it falls slowly.

This section highlights that under some circumstances, assortativity could mechanically rise even when the level of homophilous preferences stays constant. The condition for this rise is when same-type preferences are strong and largely unexpressed because of groups imbalances. Thus when a system is at the beginning of a process of social transformations which change its demographic composition under-

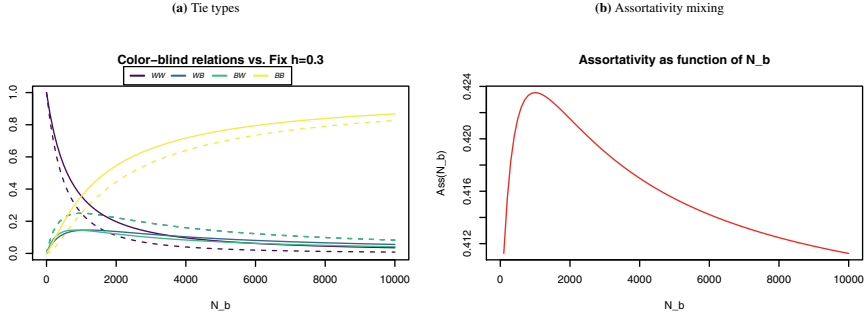
---

sample. However, I consider this a possible limitation that might be context specific and it needs further research and data.

<sup>38</sup>For simplicity students and professors are added with equal sizes, yet the main result here is consistent with a case with different rates of sub-population growth.

<sup>39</sup> $N_b$  is equal to the number of black students plus black supervisors where both have equal size.

**Figure 2.6:** Model predictions as a function of number of black in the population with a starting population of 1000 white. The model is with one population ( $S_w=T_w$ ) and  $h=0.3$  for all groups. Solid lines for model with  $h=0.3$  and dashed lines for model  $h=0$ .



lining preferences will become apparent thanks to size changes of group interaction. Therefore, I want to remark here that the increasing tendency of homophily found in section 2.8 could be partially artificially due to the increasing number of black who entered the system. It suggests that social transformations in South Africa academia may still be in the transition phase. These findings suggest an additional complication in the study of homophily in a context of social transformations. In such a context, standard measures are open to misleading interpretations. Paradoxically, on the one hand, apparent changes in preferences might, in fact, be due to changes in the relative sizes of the demographic composition of a population and the interaction of thereof. On the other hand, this suggests that a process of inclusion may in the initial phase creates tensions due to rising levels of choice homophily.<sup>40</sup>

## 2.11 Conclusion

Since the end of apartheid in 1994, South Africa is changing and slowly de-segregating from a society divided on racial lines by the law to a more integrated society.

Social transformations are ongoing; in many walks of life still exists a considerable gap between aspirations and the current situation. The demographic composition of formerly white universities is changing, yet the process is unevenly distributed in the system. How effectively and quickly transformations proceed depends on how the relations among students and supervisors evolve across race and gender lines.

I have examined the composition of tie formation in the relations between the students and their thesis advisors in South African universities. Following the end

<sup>40</sup>This has a similar implication to those of conflict theory that is described in the review of the literature.

of apartheid and the formal de-segregation of the universities, the data underlines an ongoing transformation in its transition phase.

I found evidence of both choice and induced homophily along racial and gender lines in the formation of student-supervisor ties. My findings have policy implications. In particular, it should be noticed that homophily can make a transition phase of social transformations, difficult, and very lengthy. Thus, the results suggest that a policy with a joint target in reducing induced and choice homophily could accelerate the transition phase of racial and gender inclusion.

Further, looking at the interaction between the demographic groups, I provide evidence that the racial(gender)-base homophily in the system might originate from white (male) students. This group might be less capable of understanding the potential gains that an inclusion policy brings for all and feel threatened.

Moreover, strong homophily in white (male) students could retain the transition phase of social transformations by hindering the creation of cross-type ties. Further, examining homophily in a context of social transformations, the integration process shows an additional complication. Paradoxically, I show that during a transition phase of social transformations homophily can mechanically increase due to changes in the demographic composition of the sub-populations.

At universities, the promotion and encouragement of cross-type supervision and mentoring (across departments and universities) could be a channel for social transformations. A slow path of social transformations is a great economic and social loss of many societies. The chapter underlines that a dual-target on induced and choice homophily in student supervision may speed up this path. In addition, such a target offers a more holistic view beyond the focus of only demographic changes and/or labor market outcomes. Given the potential tensions of a long transition phase monitoring and targeting induced and choice homophily could create a better and more productive environment for all.

## Chapter 3

# Doctoral productivity and the role of gender and race in supervision

The chapter explores gender differences in productivity of Ph.Ds graduated between 2000 and 2014 in STEM in an emerging economy with the recent systematic introduction of doctoral programs, South Africa.<sup>1</sup> I study whether student-advisor gender couple matters for student publication productivity. The chapter explores gender couples in the whole sample and it investigates the joint effect of gender and race looking at the sub-samples of (1) white-white; (2) black-black; and (3) black-white student-advisors supervisions. Further it asks whether differences among gender couples change for high, medium and low student “productivity type”. The results of this chapter show early career productivity differences: female students on average produce 11%-22% fewer article than male and this in particular for female students working with a male advisor and not those working with a female one. Further, the difference in publication productivity with the male-male student-advisor couple is u-shaped across student productivity. In particular, female students working with female advisors with low and high productivity do not differ to male-male couple. When

---

<sup>0</sup>This chapter is available also as Rossello et al. (2020) in the UNU-MERIT working paper series. Financial support was provided through the Institut Universitaire de France. I gratefully acknowledge the comments and suggestions of participants of the workshop “APEX – Academic Pipeline Development for Excellence” at CREST-SciSTIP as well as those of Dr. Gloria Bernal, Dr. Arjan Schakel, Prof. Louis Volante, Prof. Kristof de Witte, and Dr. Tatenda Zimyemba.

<sup>1</sup>The notion of the systematic introduction of doctoral programs refers to the fact that some South African professors do not have a Ph.D. in the period of the analysis. This is a feature that differentiate the South African context to the US and many European countries.

gender is couple with race the gap interests only the top 10%-20% of the more productive students. Interestingly, male students working with female advisor display no significance difference in productivity compared to male students working with male advisors. Overall, cross racial supervisions displays less severe gender differences. The findings underline that gender gap do not interest the whole population and that female advisors and cross-racial ties might have an important role for students and to reduce gender and racial inequalities.

## 3.1 Introduction

The gender gap in publishing is well documented: depending on the context, discipline, geography or era, female scientists are found to produce fewer papers per year than their male colleagues (Allison and Stewart, 1974; Cole and Zuckerman, 1984; Fox and Faver, 1985; Mairesse and Pezzoni, 2015; Holman et al., 2018; Lerchenmueller and Sorenson, 2018; Mairesse et al., 2019; Pezzoni et al., 2016).

The ultimate sources of this gap remain elusive, though Mairesse and Pezzoni (2015) have found that when biases in promotion decisions, and the frequency of “idle periods” are controlled for, women are in fact more productive than men.<sup>2</sup> They admit, though, that their context is specific, and they do not claim to have presented the universal explanation.

It is common to observe in studies of the gender gap that age plays a role in publishing productivity and that the gap can change with age (Kelchtermans and Veugelers, 2011). This observation, combined with the well-known Matthew Effect (Merton, 1988) suggest that productivity gaps might originate very early in the career. An important open issue then is whether publishing productivity gaps are observed early in the career (David, 1993; Conti and Visentin, 2015b), and if so how to understand them. The chapter can get at this issue by examining publication of scientists during the course of their doctorates.

While the study of the gender gap focuses on single scientists, it must be acknowledged that much publishing involves more than the focal author (Wager et al., 2015; Chuang and Ho, 2014; Larivière, 2012). Not only co-authors, but research assistants, co-workers, technicians, conference participants, and many others contribute with work, ideas, and suggestions. Of course, when Ph.D. students are considered as a (co-)author, the thesis supervisor is very likely to provide important input.

---

<sup>2</sup>The context of their study is France 1982-2005 and they look at 2811 scholars in Physics in universities and Centre national de la recherche scientifique (CNRS). A similar study has been done in Mexico and South Africa, and finds that, after controlling for promotion biases, female are 8% more productive than male and that there are no differences in terms of publication quality (Rivera León et al., 2017)

Often the thesis advisor is the first person with whom a student co-authors, but additionally, supervisors play a key role in introducing students into the profession. It seems very likely that the properties of the supervisor matter for a student's early success (Li et al., 2019). A priori, there are some obvious traits of the supervisor that will matter: extent of supervision, publishing record, status in the profession, quality, and so on. But other literature suggests gender (race) might also matter. For example subtle gender and racial biases can distort the meritocratic evaluation of the students. An experiment in a sample of 127 biology, chemistry, and physics professors in the USA, asks academics to evaluate the CV of students for a laboratory manager position, where gender was randomly assigned to CVs. It finds that both male and female faculty judge female students as less competent, less likely to be hired than an identical male student, and also offered her a smaller salary and less mentoring (Moss-Racusin et al., 2016). Such biases can also reduce a student's access to relevant information. A similar randomization experiment finds that black students are less likely to receive warning information from academic advisors than are white students, when race is randomly assigned to student academic records (Crosby and Monin, 2007).

Gender (or racial) bias can play a role through both sides of the relationship (Rossello and Cowan, 2019). From the student side, in education and learning the gender of the advisor can affect performance and beliefs (Gaule and Piacentini, 2018; Breda et al., 2018; Rossello and Cowan, 2019). For example, female role models are often found to be more effective in inspiring female students (Bettinger and Long, 2005; Lockwood, 2006; Aguinis et al., 2018). A recent French experiment among senior high school students finds a reduction of stereotypes associated with jobs in science, after students were exposed to a female scientist (Breda et al., 2018). In the same study, enrolment in a selective science programme increased by 30% among the higher achieving students. And in particular, the share of female (male) students in STEM programs were 38% (28%) than that in classes that did not receive the intervention.

Thus, one might expect to see female students performing better with female advisors. In South African academia, after controlling for preferential attachment and institutional constraints, Rossello and Cowan (2019) find preferences for same gender (race) in student-advisor tie formation in a sample of bachelor, master and PhD students and advisors based on enrolment data. In particular male (white) students have high tendency to form same-gender (race) relations, while among professors it is female (black) faculty who display the higher frequency.

From the advisor side, the gender of the student can also be relevant. Each Ph.D. student shares with others a thesis advisor who guides and supervises the research, provides access to knowledge (tacit in particular), co-authors, resources, and job opportunities. Thus, gender biases in this phase can limit the access of the student to resources and information. Past research has found that supervisors provide more psychological support to protégés of the same gender (Koberg et al., 1998; Aguinis et al., 2018); male advisors were more likely to agree to a mentoring meeting with a male student than with a female student with same characteristics (Milkman et al., 2015); and less willing to supervise female students (Moss-Racusin et al., 2016).

Exploring the relationship between gender (race) and performance in the student-supervisor pair is a step towards understand productivity differences among different groups within academia. Past research in Science, Technology, Engineering and Mathematics (STEM) is available only for the US in a first-tier institution (Pezzoni et al., 2016) or in a single field (Gaule and Piacentini, 2018). Looking 20,000 Ph.D. graduated between 1999 and 2008 in US chemistry departments, Gaule and Piacentini (2018) find that same-gender couples tend to be more productive during the Ph.D., and that female students working with female advisors are more likely to become faculty members compared with female students working with a male advisor. In contrast, Pezzoni et al. (2016) study all fields in STEM with data based on 933 Ph.D. graduates and 204 advisors at the California Institute of Technology (Caltech) between 2004 and 2009. In terms of student publication productivity, they find no difference between the female-female and the male-male couples. However, they find that male students working with a female advisor perform better than male-male peers, while female students working with male supervisors perform worse than male students working with male advisors.

In this chapter, I study whether Ph.D. students' early career productivity is affected by the gender (and/or race) of student and supervisor. In the first analysis I ask simply whether there is a correlation between students' publication output and the gender (race) of the student, and, independently, the gender (race) of the supervisor. In the second analysis I ask whether there is an observable effect of the student-supervisor pair. The data are drawn from an emerging economy, namely South Africa, where resource constraints in the science system generally, and universities in particular, are much more severe than they are in developed countries. One might expect that in the presence of resources constraints the "privileged group" will have better access and therefore higher productivity relative to others. The academic science system in

South Africa is relatively small — in 2012 there were only 2174 full professors,<sup>3</sup> and the production of Ph.D.s is concentrated in a relatively small number of institutions (Cowan and Rossello, 2018).<sup>4</sup> These features are typical of many developing countries (Nchinda, 2002; Gonzalez-Sauri and Rossello, 2019).

While concerns with gender in science are common to many countries today, given the history of apartheid and its on-going legacy, in South Africa there is a second axis of concern, namely race.<sup>5</sup> In South Africa people of colour make up 90% of the population and apartheid essentially excluded them from academia. Until 1994 there were “black universities”, but they were severely under-funded and not expected to do any meaningful research. One of the ongoing efforts of governments since 1994 has been the transformation of the university system to include more of the black population in the “top” (formerly white) universities. Part of the challenge has been the academic “pipeline”: whether or not faculties are trying to hire formerly excluded groups, if there is no supply of them, the system will not transform from a white male bastion to a more inclusive institution. Given that academic appointments are often heavily based on performance during graduate studies, understanding gender and race effects on PhD student publishing becomes something of significant importance in this context.

Pezzoni et al. (2016) have done a similar analysis using data from the Caltech, an elite institution in US, and this chapter is modelled on theirs. They found that compared to the male-male student-advisor couple: female students working with male advisors publish 8.5% less; and male students working with female advisors publish 10% more. Their data were constrained to a single, rather specific (in terms of student and faculty quality, and finances, just to mention two dimensions) institution, namely Caltech. Mine involves the entire national academic science system, and so might be considered more representative of national trends and effects. In addition, the statistical analysis differs from theirs in an important respect.

They study the relation between the student-advisor gender couple and productivity employing a model of cumulative advantage in science, where future productivity is explained by past productivity. They use for the estimation OLS panel regressions. In contrast, I use a quantile regression approach. The advantage of this approach is that it does not assume that the discrepancies between genders (or races) are constant throughout the population being studied. It could be, for example, that the average

<sup>3</sup>These data are available at <https://africacheck.org/reports/how-many-professors-are-there-in-sa/> last access November 2019.

<sup>4</sup>For a further discussion on the South African system see Rossello and Cowan (2019), and the report Mouton et al. (2015)

<sup>5</sup>“Race” is sometimes considered a contentious concept (and word) but in the context of South Africa it is well understood as central to the construct of the society, so I will employ the word and concept here in the way it is done in South Africa.



gender gap is driven by a few outliers in one or other of the groups. More specifically, differences in gender-specific productivity may change depending on the “productivity type” (high, medium or low) of the student. These sorts of differential effects across different population groups effects are more likely to be important where the output variable is skewed and has a fat tail, as is the case with publications. (Petscher and Logan, 2014).

Further, the sample is characterised by a close to balance population of students in terms of gender and race and allows to explore both dimensions. In the gender (race) analysis to study the gender (race) couple of student and advisors I look at the sub-samples of white (male) students working with white (male) advisor, black (female) students working with black (female) advisor and black (female) students working with white (male) advisors.

To preview in the results:

The regression analysis shows that female students on average produce 11%-22% fewer articles than males. However, dividing the two student sub-populations by student-advisor gender, and controlling for race, I find that this average gap is mostly driven by female students working with male advisors; the gap does not seem to exist for female students working with female advisors once I consider the joint effect of gender and race.

For the whole sample, the quantile regression shows that this gender gap might be non-linear and indeed u-shaped over student productivity. That is, the gender publication gap is zero among low and high publishers; but in the middle of the publication distribution female students publish less than males. But decomposing the population over gender and race, the gap involves only the  $e$  most productive students (the top 10%-20% of the publication distribution) and is strongest in black-black supervision<sup>6</sup> and not present in the population of black students working with white advisors. Interestingly, male students working with female advisors seem to have roughly the same productivity as male students working with male advisors.

Similarly, I compare black and white students and I do not find a significance difference in productivity between the two.

The findings imply first that the gender and race composition of the student-supervisor couple matters, but that the gender gap does not pertain to the whole population of scholars. In particular, among the more productive, results suggest a positive role of female advisors and that female students working with male advisors is where differences in productivity are largest. The results show that looking only at aggregated data may produce quite mis-leading conclusions, and therefrom possibly

---

<sup>6</sup>I have to remark that in this sub-sample female advisors are few.

inappropriate or ineffective policy.

## 3.2 Material

### 3.2.1 Data

The data originate from the National Research Foundation (NRF) database of South African Academia.<sup>7</sup> The NRF has a system, in which academics at South African universities apply to be “rated”. This rating has (until recently) financial and prestige incentives, so most academics in South Africa who pursue a research career do apply. Overall, rated scholars comprise about 30% of South African scholars who produce roughly the 90% of the country peer reviewed output. The STEM fields have been part of the system longer than have SSH fields, and in these fields coverage appears to be more complete. Consequently I restrict attention to this group, where the agency has a primary role in funding research. I create a unique dataset using data supplied in the application process. The raw data include student PhD supervision from 2000 to 2014, and publications from 1961 to 2014.

From student and supervisor characteristics I am able to match to NRF publication data. To be confident that publication data are complete, I include in the analysis only Ph.D. students in STEM who became active scholars in the NRF system. This constitutes 25% of the total PhD graduates over the period. The final sample represents Ph.D. students within the enrolment period of 2000-2012 and with a graduation period up to 2014. In the sample Ph.D. students graduate on average after 3.8 years after enrolment (with a median of 4 years and a maximum of 12 years). To construct the panel data I carry forward for each individual student the time from enrolment year up to two years after graduation obtaining a total of 6049 observations representing 924 Ph.D.s and 549 thesis supervisors.

In the period 2000-2014 the number of Ph.Ds graduated increased rapidly.<sup>8</sup> The sample, in table 5 in the appendix, shows similar trends and has a good representation in terms of the distribution of Ph.D. graduation over time. However, the last two periods have lower number of graduates relative to national statistics. This is because

---

<sup>7</sup>NRF is a state agency that has as its mission the promotion of research and the development of national research capacity. <https://www.nrf.ac.za/>

<sup>8</sup>As reported in the data of the Council of Higher Education (CHE) available for the period 2008-2012 at [https://www.che.ac.za/focus\\_areas/higher\\_education\\_data/2008/graduates](https://www.che.ac.za/focus_areas/higher_education_data/2008/graduates); [https://www.che.ac.za/focus\\_areas/higher\\_education\\_data/2009/graduates](https://www.che.ac.za/focus_areas/higher_education_data/2009/graduates); [https://www.che.ac.za/focus\\_areas/higher\\_education\\_data/2010/graduates](https://www.che.ac.za/focus_areas/higher_education_data/2010/graduates) [https://www.che.ac.za/focus\\_areas/higher\\_education\\_data/2011/graduates](https://www.che.ac.za/focus_areas/higher_education_data/2011/graduates); [https://www.che.ac.za/focus\\_areas/higher\\_education\\_data/2012/graduates](https://www.che.ac.za/focus_areas/higher_education_data/2012/graduates).

**Table 3.1:** Students and Advisors, by Race and Gender. A professor can supervise more than one student.

	Advisor					
	White Male	White Female	White	Black Male	Black Female	Black
Stud. White Male	179	53	232	13	4	17
	55%	37%	49%	54%	36%	49%
Stud. White Female	149	92	241	11	7	18
	45%	63%	51%	46%	64%	51%
Stud. White	328	145	473	24	11	35
	100%	100%	100%	100%	100%	100%
Stud. Black Male	123	45	168	97	17	114
	71%	64%	69%	64%	81%	66%
Stud. Black Female	51	25	76	54	4	58
	29%	36%	31%	36%	19%	34%
Stud. Black	174	70	244	151	21	172
	100%	100%	100%	100%	100%	100%

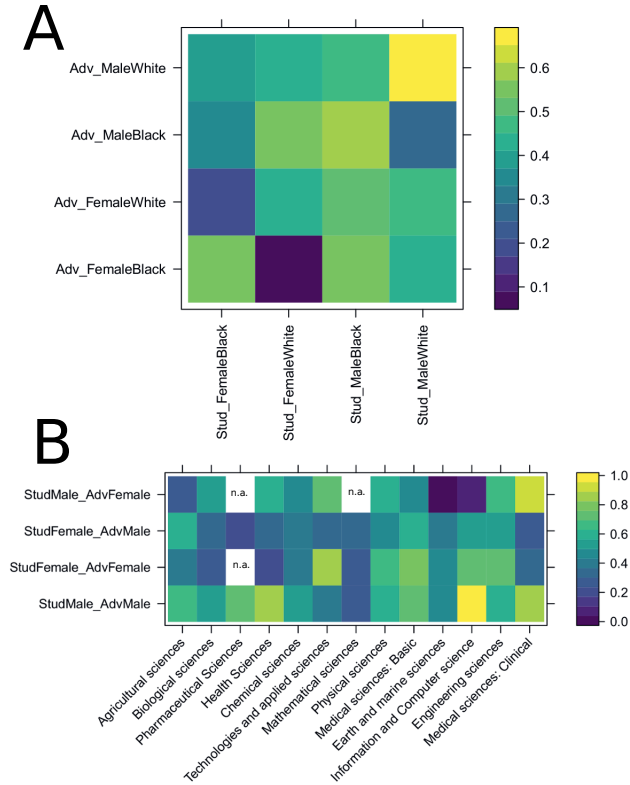
it takes several years after graduation before a faculty member is ready to apply for rating. So by restricting to students who eventually do apply for rating, I will under-sample the later years. Distributions of Ph.D.s graduated over disciplines, in table 6 (appendix B.6), are also in line with national statistics.<sup>9</sup>

Students in the sample are 58% (249 white and 282 black) male and 42% female (259 white and 134 black). Professors in the final sample are 73% male (298 white and 104 black) and 27% female (130 white and 17 black).<sup>10</sup> Table 3.1 shows the population composition in terms of student and advisor pairs. The majority of students are supervised by white male advisors (54%) followed by white female (23%), black male (19%), and black female advisors (3%).

<sup>9</sup>As reported in the data of the Council of Higher Education (CHE) available for the period 2008-2012 at [https://www.che.ac.za/focus\\_areas/higher\\_education\\_data/2008/graduates](https://www.che.ac.za/focus_areas/higher_education_data/2008/graduates); [https://www.che.ac.za/focus\\_areas/higher\\_education\\_data/2009/graduates](https://www.che.ac.za/focus_areas/higher_education_data/2009/graduates); [https://www.che.ac.za/focus\\_areas/higher\\_education\\_data/2010/graduates](https://www.che.ac.za/focus_areas/higher_education_data/2010/graduates) [https://www.che.ac.za/focus\\_areas/higher\\_education\\_data/2011/graduates](https://www.che.ac.za/focus_areas/higher_education_data/2011/graduates); [https://www.che.ac.za/focus\\_areas/higher\\_education\\_data/2012/graduates](https://www.che.ac.za/focus_areas/higher_education_data/2012/graduates).

<sup>10</sup>The sample demographic composition is close to that of the system in that period. See Rossello and Cowan (2019) for further discussion.

### 3.2.2 Variable description



**Figure 3.1:** Heat-map of doctoral average annual productivity for student and advisor gender (racial) combinations. The color intensity of each entry represents the average annual productivity of each group. Darker (lighter) colors represent lower (higher) productivity values. Productivity is  $\log(1 + pub_t)$ , where  $pub_t$  is number of student publications between year  $t$  and  $t + 2$  inclusive, divided by 3. Rows in sub-figure A are advisors gender-race type while columns are student gender-race type. In sub-figure B rows are student advisor gender couples and columns are fields.

For the average student, looking at three year moving windows between enrolment and graduation plus 2 years, the annual average number of publications is close to one; where white male students have the most (1.58) followed by black males (1.37), white females (1.21) and black females (0.75). The median values are close to zero for all groups (See Table 3 in the appendix). Publication data, also referring to a 3 year average, are skewed, and 44% (410 Ph.D.s) of the students do not publish at all between enrolment years and two years after graduation.<sup>11</sup>

I define student productivity as  $\log(1 + pub_t)$  where  $pub_t$  is number of student pub-

<sup>11</sup>Details on relative publication rates over time are shown in Figure 9 (a)(b) in the appendix.

lications between year  $t$  and  $t + 2$  inclusive divided by 3. Raw differences in student productivity between different populations and student-supervisor pairs are presented in figure 3.1 and table 7 in the appendix. Figure 3.1(A) shows that same-type supervision (2nd diagonal) correlates with higher average productivity (lighter colours).<sup>12</sup> Further, looking at students (by columns) and advisors (by rows) type productivity displays a large heterogeneity across supervision couples, suggesting a complex joint effect of gender and race. I explore this further in table 7(e) in the appendix: looking at black students working with white advisors, it is the couple (black) males with (white) female advisors who publish most. Similarly in the population of female students working with male advisors (table 7(f) in the appendix) white (female) students working with black (male) advisors have the highest average productivity. Interestingly, the female students who stand out in terms of productivity, in the top decile of the productivity distribution are those who have supervisors of opposite race (Figure 12(d) in the appendix).

Figure 3.1(B) shows average productivity across student-advisor gender and disciplines. In 5 out of 13 fields the couple female student with female advisor has the highest average productivity.<sup>13</sup> In 2 fields, Mathematics and Medical: clinical, cross gender ties are those with the highest averages. In the remaining 6 fields the couple male student with male advisors has the highest average productivity. Overall, the couple female student with a male advisor have the lowest average productivity for 6 out of 13 fields.<sup>14</sup>

### 3.3 Methods

The raw data indicate that female and black Ph.D. students publish less than male and white students. But these differences could be driven by many things. In the analysis that follows I control for several factors that are likely to contribute to a scientists' publication productivity in order to isolate the effects of gender and race. The variable of interest is the number of publications produced by a student in a year during the course of the doctorate. Because this variable has a skewed distribution, I work with logs:  $\log(1 + \text{pub}_t)$ . There is always a lag between the date of (completion of) the research and publication, so I include in the definition of  $\text{pub}_t$  publications of

<sup>12</sup>White females are an exception: they display higher averages when they work with black male supervisors, however the group has very few observations.

<sup>13</sup>The 5 fields are: Technologies and applied sciences, Physical sciences, Medical science: basic, Earth and marine sciences, Engineering

<sup>14</sup>In table 8 in the appendix I check whether there are any environmental effects at the level of university or field in terms of gender and racial likelihood of supervision association. There are not any identifiable environmental effects in terms of gender in the sample. But there are along racial lines, thus I run the analysis on separate racial sub-sample of the data.

which the student was a (co-)author, between years  $t$  and  $t + 2$  inclusive. I normalize for annual output by dividing by 3.

In line with Pezzoni et al. (2016) I ask whether doctoral productivity differs with respect to: (1) student gender; (2) advisor gender; (3) the genders of the student-advisor pair. I estimate panel OLS regressions with robust clustered standard errors.<sup>15</sup>

In all regressions I control for discipline, enrolment year, time to graduation, whether the student had published previously, whether the student has more than one advisor, and advisor productivity as the log of average publications of the advisor lagged one year.<sup>16</sup>

In table 9 in appendix B.7 I control for the joint effects of gender and race exploring the interaction terms. This preliminary analysis shows that the main difference in productivity is between male and female students: race has no role. Since the end of apartheid, the progressive introduction of black was not uniform across gender. Black females are under-represented both among students and professors, particularly in STEM fields. For this reason, I also run the analysis on different sub-samples of the data to decompose the possible joint effects of gender and race.

For the gender analysis I look at the sub-samples: white students with white advisors; black student with black advisor; and black student with white advisor. Similarly, given the context of the country, I run a parallel analysis in section 3.4.1 to compare black and white. Here I look at the sub-samples: male student with male advisor; female student with female advisor; and female student with male advisor.<sup>17</sup>

As a further contribution to understanding where the gaps originate I use quantile regression to examine the effects of the student-supervisor pair, where the quantiles are defined over productivity. This permits to observe that much of the difference in average publication between male and female, white and black students is driven by differences in the right hand tail of the output distribution.

The quantile regression formulation is:

$$Q_\tau(Y_{it}|Z_i, X_{it-1}) = \alpha_\tau + \gamma_\tau Z_i + \beta_\tau X_{it-1} + \varepsilon_{it} \quad (3.1)$$

where  $Q_\tau(Y_{it}|Z_i, X_{it-1})$  is the  $\tau$ th quantile regression function,  $Z_i$  are time invariant

<sup>15</sup>As robustness check I report Poisson panel regressions in appendix B.9 with robust clustered standard errors. The results are not qualitatively different.

<sup>16</sup>In appendix B.14 I show results of the OLS panel regressions with the main variables of gender and race and controlling only for field and enrolment years.

<sup>17</sup>There are too few white students with black female advisors, and male students with black female advisors to give reliable results for those groups. Hence there were not included.

covariates and  $X_{it-1}$  are time variant lagged controls and  $\varepsilon_{it}$  is the error term.<sup>18</sup>

The models compare: student's gender, advisor's gender and the student-advisor gender couples. The main independent variables are: the dummy *StudFemale* equal to 1 for female students and zero otherwise; the dummy *AdvFemale* which is 1 for female advisors; and the dummies for the different student-advisor couples *StudFemale\_AdvFemale*, *StudFemale\_AdvMale*, *StudMale\_AdvFemale* where the baseline category is the pair male students with male advisors. For each model I show results for the whole population (ALL) and partitioning the data according to student-supervisor racial composition (WW for white students white advisors; BB for black students black advisors; BW for black students and white advisors).

---

<sup>18</sup>For the estimation I use robust clustered standard errors to account for heteroskedasticity and intra-cluster correlation as described in Machado et al. (2011)

**Table 3.2:** Pooled OLS Panel Regression with robust clustered standard error. The dependent variable is log of 1+ average productivity in terms of number of paper between the period t and t+2. Models (1) for the Students comparison; Model (2) Advisor comparison; Model (3) Couples Comparison. Where Columns (a) On the sub-sample White Student White Professors; Columns (b) On the sub-sample Black Student Black Professors; Columns (c) On the sub-sample Black Student White Professors. Additional controls are moreAdv, logprofcumavgprod, DummyStudPrevPub, timegrad, field, and enrolment year.

	(1)	(1)a	(1)b	(1)c	(2)	(2)a	(2)b	(2)c	(3)	(3)a	(3)b	(3)c
	ALL	WW	BB	BW	ALL	WW	BB	BW	ALL	WW	BB	BW
StudFemale	-0.113** (0.0357)	-0.116* (0.0497)	-0.217* (0.0847)	-0.0908 (0.0686)								
AdvFemale					-0.0393 (0.0416)	-0.0248 (0.0589)	0.0989 (0.151)	-0.0678 (0.0720)				
StudFemale_AdvFemale									-0.120* (0.0571)	-0.113 (0.0745)	-0.270 (0.239)	-0.123 (0.129)
StudFemale_AdvMale									-0.136** (0.0417)	-0.142* (0.0580)	-0.198* (0.0958)	-0.112 (0.0810)
StudMale_AdvFemale									-0.0733 (0.0568)	-0.0622 (0.0899)	0.0768 (0.172)	-0.0931 (0.0839)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	6049	3083	1099	1641	6049	3083	1099	1641	6049	3083	1099	1641
R <sup>2</sup>	0.284	0.354	0.260	0.263	0.280	0.349	0.244	0.260	0.285	0.355	0.260	0.264
Standard errors in parentheses												
* p<0.05      ** p<0.01      *** p<0.001"												



## 3.4 Results

Table 3.2 shows the results of OLS estimations of three models. Model 1 compares female with male students. Results for the whole sample (column 1) show that female students produce on average 11% fewer papers than male students. Looking at white students working with white advisors (column 1a) I find that female students produce on average 12% less than male students. Looking at black students working with black advisors (column 1b) I find that female students display a larger gap — 22% compared with male. Finally, among black students with white advisors, (column 1c) there is no difference between male and female students.<sup>19</sup>

Using the same structure, model 2 (table 3.2) compares male and female advisors. I find that female and male advisors have students that are not statistically different in productivity.<sup>20</sup>

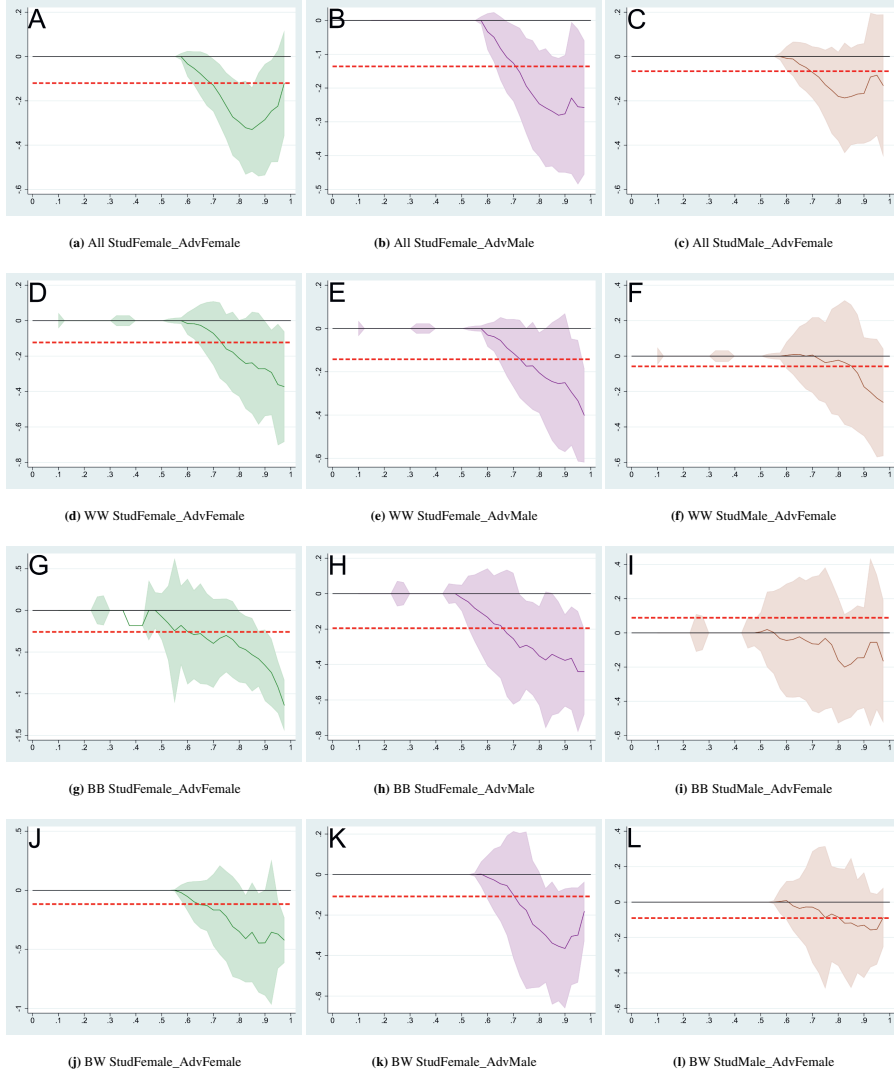
Model 3 in table 3.2 explores the gender pairs of students and advisors. In this model the baseline category is the student-supervisor pair male-male. Overall female students working with male advisors have the largest gap compared with the male-male couple: they produce on average 14% fewer papers; while it is 12% fewer for female students working with female advisors. Male students working with female advisors do not differ in productivity with male-male. Decomposing the joint effect of gender and race, the gap in productivity between female and male students is mainly driven by female students working with male advisors in same-race supervisions. In particular, I find that when student and supervisor are both white, (column 3a) female students working with male advisors produce 14% fewer papers than a male student working with male supervisor. Similarly, among black-black supervision pairs (column 3b) female students working with male advisors produce on average 20% fewer papers than do males students working with male advisors. Interestingly, the group of black students working with white advisors (column 3c) display no significant difference in productivity between gender couples.

---

<sup>19</sup>I should underline that in the BW sub-sample, outstanding students (top 10% more productive) are females and have a median productivity higher than males. However, they comprise less than 25% of their relative population (Appendix B.5 figure 12).

<sup>20</sup>It is important not to draw hasty conclusions from this result. It is consistent with a situation in which male advisors favour male students, thus having productive male and unproductive female students, and female advisors doing the reverse. This kind of homophilous preferential attention, were it to exist, would produce the results observed here. This observation should not be read as a conjectural explanation, though, but rather a caution against quickly interpreting this to mean that advisors are gender-blind.

**Figure 3.2:** Quantile Regressions for student annual average doctoral productivity comparing student-advisor gender couples. Productivity is  $\log(1 + pub_t)$ , where  $pub_t$  is number of student publications between year  $t$  and  $t + 2$  inclusive, divided by 3. Each row shows results for a different data sample: All (A, B, C); only white student-advisor (D, E, F); only black student-advisor (G, H, I); and black student with white advisor (J, K, L). In each sub-figure, the horizontal axis represents percentiles and the vertical axis shows estimated productivity difference of student-advisor gender couple with the baseline Male-Male couple. The columns show respectively estimated coefficients for productivity difference for the dummy female-female (green), female-male (violet) and male-male (brown) student-advisor couple. Quantile regressions are done for each 2.5 percentile using robust clustered standard errors according to Machado et al. (2011) and estimates for the student-advisor gender are shown with 95% confidence intervals. The solid black line is zero, dashed red line is the (non-quantile) panel OLS estimation of Models 3 from table 3.2. Additional controls are: discipline, enrolment year, year, time to graduation, whether the student had published previously, whether the student have more than one advisor, the log of average publications of the advisor lagged one year. Corresponding regression tables are in Appendix section B.10



To go beyond average differences and to accommodate the skewness and fat tails of the dependent variable, I explore model (3) using quantile regressions with clustered standard errors. In this way I am able to look for the origin of this difference, and ask whether discrepancy between groups is stronger or weaker for different parts

of the population, where the population is sorted into quantiles by publication productivity.

Figure 3.2 presents quantile regression estimates done over 40 groups (each representing 2.5% percentile of the population) for whole data and sub-samples. It shows the coefficients of the dummies student-supervisor gender pairs with 95% confidence intervals where zero represents the male-male supervision baseline. Results for the whole sample show that productivity differences of female-female and female-male with the male-male couple is u-shaped over student productivity (fig.3.2(a)(b)(c)). The u-shaped productivity gap is most pronounced for female students working with female advisors, who are not statistically different from male-male for low (<70th percentiles) and high (>90th percentiles) student productivity type (fig.3.2(a)). Female students working with male advisors overall display larger gaps with the male-male couple in line with OLS results (fig.3.2(b)).

Results for data sub-samples look at the joint effect of gender and race. Overall I find that the productivity gap (with the baseline male-male couples) increases with student productivity. The figure shows that female students working with female (fig.3.2(d)(g)) or male advisors (fig.(3.2(e)(h)(k))) compared to males working with males occurs mostly after the 75th percentile of the productivity distribution and tends to grow with publication productivity. In line with OLS results, male students working with female advisors are not different in productivity compared to male-male pairs.<sup>21</sup>

I explore this evidence further in appendix B.10. For the sub-sample of white students working with white advisors, table 15 shows that differences in productivity of female students working with female advisors compared to male-male exists only among the most productive — top 10%, 5%, 1% (90th, 95th, 99th percentiles) of the students and ranges from a 27% to a 41% difference. I find a more heterogeneous and pronounced difference (from -20% to -47%) for the female students working with male advisors. The difference is significant also for the top 20% (80th percentile) productive students.

I find similar results in the sub-sample of black students working with black advisors (table 16) and with white advisors (table 17).

Results of the quantile regressions run on the entire population underline that the productivity difference between male and female students is most pronounced for females working with male advisors and u-shaped, in particular for females working with a female advisor. However, when student-supervisor gender is coupled with race the gap with male-male is not u-shaped but rather downward sloping and in fact ex-

<sup>21</sup>However, this group displays significantly lower productivity than male-male at the 99th percentile for the sub-sample of white-white and black-black supervisions (see table 15 and 16 in appendix B.10).

ists only among the top productive (top 10-20%) students especially for same-race couples. This suggests some composition effect reminiscent of the Simpson paradox. The Simpson paradox underlines that aggregate figures can show opposite trends to disaggregate ones. Indeed, one of the well-known instances of the paradox concerns gender or racial sorting into scientific disciplines and universities (Mullen and Baker, 2008). I test for such environmental effects at the level of university and field in table 8 in Appendix B.5, there were not any identifiable effects looking at gender but I found some along racial lines. This may explain differences in the regression results for different sub-populations.

### 3.4.1 Results looking at race

I perform the same analysis comparing black and white students and advisors for the whole sample and 3 sub-samples of gender couples: male-male (MM), female-female (FF), female students male advisors (FM). In appendix B.11 and appendix B.13 I show respectively results for OLS and quantile regression estimation. The results show no difference in productivity between white and black students for all sub-samples. This is particularly relevant for policy in South Africa Academia. This year in South Africa many Ph.D. funding schemes (in social sciences in particular) are ending and they will be re-discussed.<sup>22</sup> Funded Ph.D.s programs are essential in a country with large inequalities like South Africa. Surveys underline that black students identify financial constraints as the main reason preventing them from pursuing postgraduate education (Mouton et al., 2015). Thus, in the South African context where financial constraints are removed to a great extent (doctoral programs in STEM are usually funded) (Mouton et al., 2015) the fact that I do not find any difference between black and white students may underline the importance of such funding schemes that guarantee access to postgraduate education for all.

## 3.5 Conclusion

In STEM subjects, in South Africa, I find some evidence of an early career productivity differences between male and female Ph.D. students: on average male students are more productive in terms of publishing during their doctoral studies. Disaggregating the population according to the gender composition of the student-supervisor pair, I find that this difference is mostly attributable to female students working with male advisors. Female students with female advisors have publication records very similar

<sup>22</sup>The NIHSS-SAHUDA funding program for example ends in 2020; available at <https://www.nihs.ac.za/content/nihss-sahuda-programme>. Last access December 2019.

to male students. Disaggregating further, and looking at the joint effects of gender and race, I find larger gaps for female students when student and supervisor are of the same race. Again, female students with male advisors show the lowest productivity. There are two striking observations about these results however. First is that most of the gap in average performance is driven by a gap in the right hand tails. That is, in the “moderately productive” group of students, males and female have very similar numbers of publications per year. It is only when I look at the very highly productive (top 20%), that a large male-female gap appears. Second, looking only at aggregated data, the relationship between the gap and productivity levels is U-shaped: among high and low productivity students men and women look very similar, but among moderately productive students the women lag. But this U-shape disappears almost completely when the student population is disaggregated. Among any sub-population if the gap is striking, it is only so among the most productive part of the student population. One very general and important conclusion is that in thinking about the issue of gender or racial gaps (whether productivity or anything else) it is very important to get the level of aggregation correct.

The results may suggest that female students with male advisors may receive lower quality (or quantity) supervision than do male students. There may be other reasons behind this difference though. The chapter do not attempt to give an ultimate explanation of such observed differences but I consider the following important lines of future research. The observed differences in early career productivity could be due to student-supervisor personal relations; access to resources; differences in the career paths; different nature of the research output in terms of content (for example between basic or applied research which can translate into differential ‘publishability’).

The personal relations hypothesis is compatible with results in Rossello and Cowan (2019), which finds a same-gender (same-race) bias in supervision-tie formation, driven largely by male (white) students and female (black) advisors. Bias in tie formation relates with group behaviour and socialisation in the working environment which may disadvantage female students working with males (Blackburn et al., 1981; Van den Brink and Benschop, 2014; Zinovyeva and Bagues, 2015). More in general, social relations are embedded in networks which are found to vary with gender and enhance or restrict access to resources, information, and collaborations (Jadidi et al., 2018).

Differences in productivity are often explained by differences in career paths induced by motherhood (see Pezzoni et al. (2016)). Past research has found that female productivity has a negative shock during the first 3 years of a newborn (Mairesse et al., 2019). In South Africa female fertility rates peak at age 25-29 which corresponds to doctoral years (Lehohla, 2015). Such a shock may be accommodated differently depending on whether the female student works with a male or with a female supervisor.

A further explanation, explaining why the female-male couple has the lowest productivity, can relate with the two-world hypothesis. This hypothesis states that there exists a gender or racial specialization in specific (sub-)disciplines (Moore et al., 2018). Thus, cross-gender couples may re-combine different (sub-)fields and knowledge. More in general, the management literature has found that diversity is associated with novelty and innovation because it is more likely to recombine distant knowledge and expertise (Rzhetsky et al., 2015; Shi et al., 2015; Uzzi et al., 2013; Chen et al., 2009; Fleming, 2001). In science novelty is often a risk, particularly for a younger scientist, and may have slower returns (Wang et al., 2017; Boudreau et al., 2016; Verhoeven et al., 2016; Azoulay et al., 2011). Taking risks early in the career may slow down productivity in the short-run affecting ‘publishability’ of the research. Different gender composition pairs may differently mitigate the risk.

All these mechanisms may individually and jointly explain the findings and contribute to an early career gender gap in science. Beyond possible causal explanations, the chapter testifies a gender productivity gap and shows how the promotion of cross-racial ties could reduce it. Indeed, the chapter underlines the interlinked nature of gender and racial ties and productivity during the training phase preceding the first academic job.



## Chapter 4

# Emergent structures in faculty hiring networks, and the effects of early-career mobility on academic performance

This chapter examines the transition into the first academic job. In particular, the chapter is about the South African job market for PhDs. PhD to first job mobility involves the preferences of both the hiring institution and the candidate. Both want to make the best choice and here institutional prestige plays a crucial role. A university's prestige is an emergent property of hiring interactions, so the chapter uses a network perspective to measure it. Using this emergent ordering, I compare the subsequent scientific performance of scholars with different changes in the prestige hierarchy. The chapter asks how movements between universities of different prestige from PhD to first job correlates with academic performance. I use data of South African scholars from 1970 to 2012 and I find that those who make large movements in terms of prestige have lower research ratings than those who do not. Further, looking only those with large prestige movements, those with higher prestige PhDs or first jobs have higher research ratings throughout their careers.

---

<sup>0</sup>This chapter has been published as Cowan and Rossello (2018), in the journal *Scientometrics*.



## 4.1 Introduction

Placement in an academic position directly after completing a PhD is one of the most stressful events in an academic career. A good first job provides access to good colleagues, and a good affiliation from which to apply for funding. Anecdotes say that an academic's reputation is made in the first 6 years after the PhD, and these anecdotes are consistent with Robert Merton's Matthew Effect — early success provides the springboard for later success in academia (and many other venues). At the same time, from the side of the institution, a “good hire” constitutes an improvement to the quality of faculty, and future improvements in reputation. Many resources are spent on both sides of the PhD job market, as recruiters and recruitees try to do as well as they possibly can.

But young faculty hiring is a classic problem of asymmetric information (Connelly et al., 2011). PhDs usually have only a thin record of citations and publications, which it means that their intrinsic quality is largely unobservable by any hiring committee. In this type of situation, a committee will look for signals of quality, one of which is the status of the university granting the PhD (Clauset et al., 2015). Moreover, since PhD to first job mobility involves the preferences of both the hiring institution and the candidate (Barnard et al., 2016; Conti and Visentin, 2015a), and both want to make the “best choice”, hiring decisions are pairwise assessments of quality between the two agents. The sorting of PhD graduates through the first-job market can be seen to imply an emergent prestige ordering of universities, encoding the collective assessment of each others' quality (Clauset et al., 2015).

This chapter is about the South African job market for PhDs. In particular, I ask whether PhD-to-first job mobility is correlated with future research performance. It looks at the South African PhD job market as a system where universities' prestige plays a role not only in hiring but may also correlate with individuals' later academic performance. In particular I ask how movements between universities of different prestige from PhD to first job correlates with future academic performance.

In the analysis I first develop a new measure of prestige of the South African universities, based on the idea that the PhD job market contains information about how universities judge each other's graduates, and so, by implication, how they view each other's quality. There can be two reasons I might expect to observe a correlation between first job placement and future academic success: If recruiters are able to identify talent, even noisily, and academics want to work at “the best” institutions, then the quality of institutions at which graduates are hired correlates with their own intrinsic quality. In this way job placement is (perhaps just) a signal of quality, and so job placement should predict future success. On the other hand, even supposing

that all graduates are of exactly equal quality, being affiliated with a top institution should give better access to resources and funds, and perhaps colleagues, which should provide a career advantage. Starting here I ask whether PhD to first job movements are in fact correlated with future academic performance.

This analysis is aimed at increasing the understanding of the university system, looking at social inequalities, and career trajectories. These issues are of particular interest in the South African context. The country is still struggling to achieve social transformation post apartheid, especially within the university system (Barnard et al., 2016). A part of the difficulty of this transformation lies in bottlenecks in the general university hiring process. There is a large black population of students and faculty in the system as a whole, but the strongest universities, or those with the strongest research reputations, are the formerly white universities.<sup>1</sup> Here blacks are badly under-represented, both in the student population and especially in the faculty. The challenge of transforming the racial composition of the faculty of the those universities with the strongest reputations is to a great extent one of hiring, and in particular of hiring recent PhD graduates. Studying the processes by which people get into this profession is one of the first knowledge gaps to cover.

This chapter reveals how the transition from PhD to first job, operating within a hierarchical system made of interactions among the different institutions, has long-run effects also on scientific performance. I show that the 5 most prestigious South African universities produce more than 50% of PhDs in the country and they tend to hire their own or each other's graduates. These findings are in line with previous US-based work which finds that faculty hiring obeys a hierarchical structure based on institutional prestige, which in turn produces or maintains social inequalities (Burris, 2004; Clauset et al., 2015). A simple Kolmogorov-Smirnov (KS) test of the distributions of prestige rank-change indicates that under-represented groups (women and blacks) are more likely to move down in prestige than are white males, when moving from PhD to first job, which can contribute to a different type of social inequality. The main concern here, though, is with the relationship between different prestige transitions from PhD to first job, and academic performance. In this respect I find two results: a positive role of inertia, and a positive role of prestige. The results on inertia show that scholars who make large movements in prestige tend to have lower performance than those who do not. The role of prestige is evident looking at scholars making large prestige movements: those with more prestigious PhDs or first jobs have higher future career performance.

---

<sup>1</sup>Language regarding racial or ethnic identity can be charged. As the empirical part of this chapter is about South Africa, I follow the standard terminology accepted there.

## 4.2 University prestige and young faculty hiring

To measure university prestige is not easy, in part due to the many definitions of prestige that are employed. Generally speaking prestige is associated with formal university rankings such as the Shanghai Jiao Tong Academic Ranking, or the Times Higher Education Ranking. But there are many other measures and methods that scholars have found to rank prestige of departments and universities: subjective survey based measures; output based measures; labour market based; or some combination of thereof all exist. During the 1970s and the 80s many works analysed the relationship between subjective reputational rankings based on surveys, and “objective” rankings based on research outputs and productivity (for example citations, citations per capita, number of paper published). Hagstrom (1971), for example, uses survey data of department prestige for hard sciences in the US, looking at the correlations between prestige and input/output variables of the universities. He finds that prestige correlates with size, research output, research facilities and opportunities, quality of faculty background, number of postdoctoral fellows, selectivity of the undergraduate program, and awards. Webster et al. (1991) present an extensive review of this debate looking at work published between 1965 and 1982. They collect 28 articles aimed at ranking Sociology departments in the US, and find both similarities and differences between reputational rankings and productivity rankings. In particular they underline the strong correlations between these two measures when the sample is restricted to the top departments. Additionally they highlight, as in the more recent contribution of Burris (2004), the persistence over time of department prestige, finding previous prestige to be the best predictor for current prestige no matter the level of previous productivity. Webster et al. (1991) conclude that future research on prestige rankings should incorporate the sociological stratification perspective in order to explain the link that university status has with job market placement and research performance. Some formal university rankings try to incorporate prestige using surveys. For example, QS world university ranking asks scholars, specialised in a specific field, to rank universities both in general and according to the discipline of the respondent.

Survey measures try to address the potential bias of productivity measures, as in fact the presence of few star scholars can boost research output of one department but not be representative of it as a whole (Barnett et al., 2010). But surveys suffer from a fundamental problem. It might be possible to have a reliable rank of the top universities (because everybody in the field knows more or less who they are and their relative stature) but moving down the ranking, at a certain point survey respondents are not able to perceive the differences between similar institutions. Part of the issue here has to do with localization: institutions (and the individuals in them) are much more

cognizant regarding the institutions with whom they are competing (for students, faculty, resources), thus it is unlikely that faculty in top-ranked institutions will be able to differentiate between the 100th and 110th ranked institutions. But the latter two, competing with each other, are much more likely to have knowledge about each others' strengths and weaknesses. Thus knowledge about relative rankings of universities is likely to be quite localized within the ranking itself. For this intrinsic characteristic of how the university system operates, survey based measures are likely to be unreliable, particularly below the top (or perhaps second) tier institutions.<sup>2</sup> The algorithm I use below to create a university ranking is consistent with this localized knowledge, and in fact takes advantage of it.

In sociology, institutional stratification in higher education refers to a social process that causes a hierarchical differentiation among the universities, with elite and prestigious schools on one side and lower status ones on the other (Shavit, 2007). University prestige enhances stratification, as Jung and Lee (2016) summarise, because it engages and attracts the talented experts and resources, often drawing them out of lower ranked universities. This causes structural inequalities within the higher education system. For example Mai et al. (2015) study the hiring network of PhDs in the field of communication in the US. They find that the hiring patterns follow a strict hierarchy, in line with the stratification hypothesis. They also find that institutions' ability to place their PhD graduates in other universities is particularly stratified. This supports the idea that the hiring network represents a bilateral assessment of quality among institutions because it signals an acknowledgement of the university that trained the PhD. More interestingly, it suggests that an examination of hiring patterns will reveal the consensus prestige ranking. Along the same lines Barnett et al. (2010) extract centrality measures of the faculty hiring network as a measure of the quality and prestige of doctoral education in the field of communication. Their logic is similar, driven by the idea that prestige rankings are emergent, and that PhD placement is indicative of how universities or departments view each other.

Bair (2003) studies the link between university prestige of American finance PhD programs and hiring. He finds that top ranked PhD programs in finance preserve their reputations by hiring each other's graduates or directly their own graduates. His findings are also linked to previous work, where this pattern is evident in prestigious doctoral programs in other fields: law schools, mathematics, physical sciences, social sciences, chemical, engineering, psychology, and social work (Bair and Boor, 1991; Bair and Bair, 1998).

Bedeian and Feild (1980) study the stratification hypothesis using US data from 24

---

<sup>2</sup>No doubt this explains why almost all rankings provide integer rankings only for a given number of top places, after which institutions are grouped into rather large groups (100-199; 200-299 and so on).

top graduate departments of management. They find that the academic placement in management departments is influenced by doctoral prestige (measured by a subjective survey-based measure); in particular they find a significant relationship between the prestige of scholars' PhDs and the prestige of their current positions. The article ends with two possible, opposing, explanations: Either merit is irrelevant and hiring processes rely on prestige only; or the prestige of people's PhD department is related to an unobserved variable indicating the scholar's intrinsic ability.

In this chapter I consider the prestige of a scholar's PhD institution but look also at transitions in prestige from PhD to first job, asking whether those transitions are correlated with future research performance. The implicit hypothesis here is that movements may indicate perceived quality of the individual that goes beyond using the PhD institution as a signal of quality.

With respect to the relationship between university prestige and performance, results in the literature are mixed and available only for North America. The different results are mainly due to different models, measures of prestige and output, samples analysed, field considered, and to the time span of academic career included, since this effect is likely to affect academic output differently as time passes (Miller et al., 2005). Moreover, there is a clear circular causation: the prestige of a scholar's PhD increases the likelihood of a prestigious job, and institutional affiliation may affect individual research output. Moreover it is well known that prestige of an institution is highly correlated with its research output. For this reason it is difficult to distinguish at the individual level whether those with prestigious affiliations are intrinsically more productive or if they gain cumulative advantages from their affiliation (Merton, 1968).

Williamson and Cable (2003) study the predictors of research productivity of 152 young management faculty who find their first academic jobs after the PhD in the period 1987-1995. Using a structural equation model, they find effects of supervisor's research productivity and department scholarly output (both for origin and placement departments) on early career performance. But they find no direct effect of the origin department's prestige. Prestige does appear to be important as a predictor of first job placement but there is no direct effect of PhD institution prestige on individual performance.

The lack of a direct effect of prestige on performance could be related to the correlation between department prestige and scholarly output, both of which are present in their model. Alternatively, it could simply be attributed to the measure of prestige they used, the Gourman Report, which has never disclosed its criteria or its methods.

Miller et al. (2005) analyse the predictors of a prestigious job, looking in particular at prestige of the training institution and individual research output. They use a sample of 445 Business school graduates between 1977 and 1985, two measures of

prestige (survey based and the Gourman Report measure), and as research output a composite index which considers the individual's journal publications and citations, discounted by contribution (i.e. single author, first author). Again using a structural equation model they find that the determinants of prestige of the employing institution are the prestige of the training department, which operates for many years of the career, and research output. In this line of research, the findings of Miller et al. (2005) are supported by results in other fields: Allison and Long (1987) (physics, chemistry, mathematics, and biology); Judge et al. (2004) (psychology); Baldi (1995) (sociology); Reskin (1977) (chemistry). On the other hand, though, Hurlbert and Rosenfeld (1992) looking at the field of psychology, find that prestige of later jobs is weakly affected by publications but not at all by the prestige of the PhD institution; similarly Long (1978) find for biochemistry no effects at all of either of those two variables.

Interest in the relationship between the university prestige and career outcomes, including social inequalities and allocation of talent, extends beyond the university faculty job market. It is well known in non-academic job placement: Jung and Lee (2016) examine the relationship between university prestige and subsequent wages of workers in South Korea. They find that university prestige, measured using standard university rankings, matters for job market outcomes, particularly regarding salary. Araki et al. (2016) study employee promotion in Japanese manufacturing industries, finding again a crucial role for the prestige of the universities where workers got their degrees. To measure prestige they rely on standard university rankings and find that in the early stage of a worker's career university prestige is crucial because it corresponds to the employer's *a priori* ideas about the distribution of abilities among workers. So, among young employees the likelihood of being promoted is higher for those with prestigious degrees because the employer will decide who to promote according to his prior. These results in the non-academic markets underline how important are these studies that try to shed light on the mechanisms driving the academic job market. The academic job market has of course its specificities, but it shares common elements with other job markets (especially those for highly specialized occupations).

In this chapter I explore the relationship between university prestige, young faculty hiring and individual research performance. Past research underlines how young faculty hiring follows a stratified hierarchy (Clauzet et al., 2015; Burris, 2004), and to connect to this literature I develop a network-based measure of the prestige of South African universities. In contrast to Clauzet et al. (2015), who consider scholars' current affiliation, I build the faculty hiring network looking at the very first job right after the PhD. If the job market is being used to infer universities' views of each others' prestige, it is essential to apply that inference where prestige is most relevant.

That is, when information about the candidates on the job market is least complete. Years after the PhD, presumably most academics have an established record, and good information about them is available. So it is the first job that contains the most accurate signal of how universities evaluate each others' quality. Moreover, Clauset et al. (2015) focused on 3 fields (computer science, business, and history) and they produced separated prestige rankings looking 60 top departments for each one. Their period of analysis is also limited to one year. The goal here is to investigate the entire university system of the country focusing at university level on a longer time scale. Again contrasting Clauset et al. (2015), I consider all universities together as an integrated system by defining broader fields, and I also take an historical perspective looking the first job market over more than 3 decades.

After computing the new measure of prestige, I compare the future research performance of scholars with different prestige movements in their PhD to first job transition. In particular, in order to distinguish between individual performance and possible cumulative advantages gained from their affiliations, I match people with different movement in the prestige hierarchy but same gender, ethnic group, PhD obtained year, and first job (or PhD) institution. So this chapter differentiates from past contributions giving a ranking of universities coherent with the mobility issue and predicting career success for different time spans looking not only at mobility but also the individual movements in the prestige hierarchies. As additional value added of the work, to measure individual career success I use a non-bibliometric measure based on the informed evaluation of international experts in the specific scientific field.

The chapter finds that inertia and university prestige are both positively related to a scholar's future performance. Related to inertia I find that the scholars who make large movements in prestige are considered to perform worse than those who do not. Related to prestige, instead, I find, looking those who experience large movements in prestige, that those with higher prestige PhD or first job perform better. This chapter addresses the knowledge gap regarding the relation between the role of university prestige in young faculty hiring and the subsequent individual performance of the scholars. This will increase awareness about the functioning of the higher education system in an emerging country such as South Africa and show whether it displays similarities with previous, mostly US based, work.

### 4.3 Data and variable construction

I use data from the South African National Research Foundation (NRF<sup>3</sup>) from 1970 to 2012 which contains detailed personal information of the scholars (i.e. gender, ethnic group, affiliation, career history, scientific field, and NRF rating). I restrict attention to academics who received their PhDs from 1970 to 2004 (when the major reform of the university system took place). This also permits to examine medium and long term effects on career. The main variable is the NRF “rating” for years 1983-2012, which is a measure of individuals’ academic performance. The process by which the NRF grades researchers is a very rigorous examination of a candidate’s research output. The NRF solicits half a dozen international referees to evaluate the CV and published papers of the candidate for rating. This process ends with a rating: a scientific committee evaluates the content of the referee reports and assigns the rating: one of 13 ordered categories. Strong institutional incentives imply that almost all academics with a research-oriented career apply to be rated: NRF data cover the 30 percent of scholars in the country who produce about 90 percent of all South African peer-reviewed research outputs (Barnard et al., 2012; León et al., 2016).<sup>4</sup>

The analysis focuses on scholars in the field of Science, Engineering and Technology (SET). In the field of Social Sciences and Humanities (SSH) language and schools of thought often put constraints on the PhD to first job transition. These constraints are particularly relevant and in the past have represented a strong check on academic mobility in the South African context. In the main text that follows I present only the results for SET, as the results for SSH are less reliable. However, I include a parallel presentation of SSH in the appendix C.19.

Separately for each field (SET and SSH) I construct the hiring network among the different South African institutions, based on scholars who found their first jobs in a South African university within 5 years of receiving the PhD. I then calculate the network-based measure of university prestige (*prestige ranking*) and for each individual his prestige rank-change from PhD to first job (that is, the difference between the prestige ranking of the PhD institution and that of the first job institution). In the next sections I present the details of the faculty hiring network, prestige ranking, and prestige rank-change.



PhD Institution	First Job Institution																					
	CapeTown	NelsonMandelaMet.	Witwatersrand	Pretoria	Johannesburg	FreeState	SouthAfrica	Stellenbosch	KwaZuluNatal	NorthWest	Limpopo	Rhodes	WesternCape	FortHare	WalterSisulu	Venda	CentralUnivOfTech	TshwaneUnivOfTech	VaalUnivOfTech	MonashSA	DurbanInstituteOfTech	CapePeninsulaUnivOfTech
CapeTown	107	4	5	0	2	0	2	19	1	2	1	3	4	1	1	0	0	1	0	0	0	5
NelsonMandelaMet.	133	0	1	1	0	0	1	0	0	1	1	1	1	1	1	0	0	0	0	0	0	0
Witwatersrand	4	2	83	6	2	1	2	1	8	0	1	1	0	0	0	0	0	2	0	0	0	0
Pretoria	2	1	1	116	7	3	7	6	1	6	2	0	3	3	0	1	0	6	0	0	0	0
Johannesburg	0	2	2	5	30	4	7	1	0	4	4	0	1	0	0	0	0	1	0	0	0	0
FreeState	0	1	0	1	1	55	0	8	0	2	3	0	0	0	0	0	6	0	0	0	0	0
SouthAfrica	0	1	1	1	0	0	15	2	1	0	0	0	1	1	0	1	0	0	0	0	0	0
Stellenbosch	2	1	2	4	1	6	1	73	3	2	1	0	6	1	0	0	1	0	0	1	2	0
KwaZuluNatal	4	0	10	5	0	1	2	4	65	1	2	1	1	2	5	0	0	0	0	12	0	0
NorthWest	0	0	2	3	1	2	1	2	1	52	3	0	0	0	0	0	1	2	0	0	0	0
Limpopo	0	0	0	2	0	0	0	0	0	0	7	1	0	0	0	1	0	0	0	0	0	0
Rhodes	2	2	4	3	0	1	0	0	3	1	2	19	0	1	0	0	0	0	0	0	0	0
WesternCape	2	0	1	0	0	0	0	1	0	1	0	0	8	0	0	0	0	0	0	0	1	0
OffFortHare	0	0	0	0	0	0	1	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0
WalterSisulu	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0
Venda	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0
CentralUnivOfTech	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0
TshwaneUnivOfTech	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0
VaalUnivOfTech	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0
MonashSA	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
DurbanInstituteOfTech	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
CapePeninsulaUnivOfTech	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1

**Table 4.1:** Adjacency matrix of the hiring network for the years 1970-2004 in SET, rows are PhD institutions and columns are first job institutions. Each entry represents the number for people with a PhD in university  $i$  hired as first job in university  $j$ .

### 4.3.1 Faculty hiring network

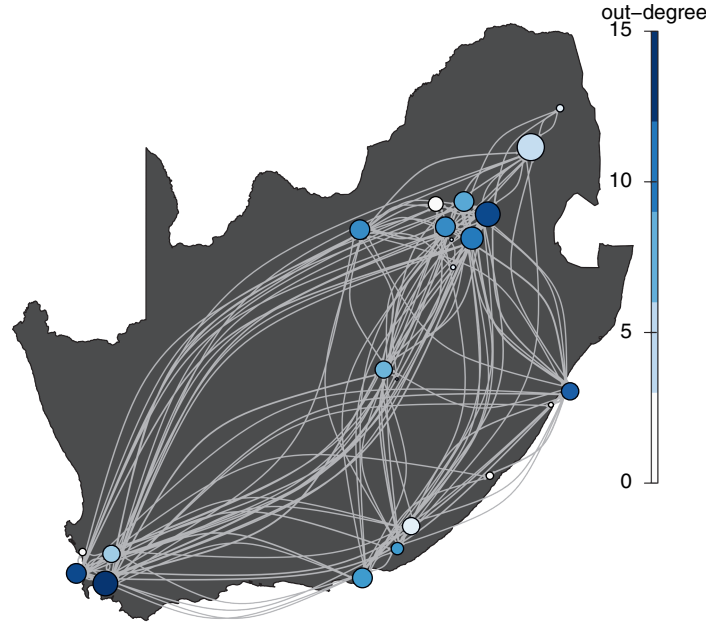
The hiring network is a weighted and directed adjacency matrix  $M$ . It has 22 rows and columns that are the 22 South African universities,<sup>5</sup> where each entry  $m_{ij}$  represents the number of scholars with a PhD from university  $i$  and a subsequent first job in university  $j$ . To illustrate, the adjacency matrix for SET is shown in Table 4.1. Summary statistics are shown in Table 4.2, with populations subdivided by gender and race.<sup>6</sup>

<sup>3</sup>NRF ([www.nrf.ac.za](http://www.nrf.ac.za)) is a state agency that has as its mission the promotion of research and the development of national research capacity.

<sup>4</sup>For more information on the rating system see <http://www.nrf.ac.za/rating>.

<sup>5</sup>In 2004 the university system was reformed: some universities were merged and changed names. The reform does not affect the big universities, excepting University of Johannesburg and KwaZuluNatal, but does have some effects on the lower ranked ones. I use the post-merger names because the data are more complete. I discuss this in the Appendix C.15.

<sup>6</sup>In South Africa there are formally four “racial” groups: black, white, Indian and coloured. The word “black” is sometimes used to refer to the aggregate of black, Indian and coloured, and this is the meaning I



**Figure 4.1:** Hiring networks 1970-2004 SET. The vertices are the South African Universities, plotted at their geographical coordinates (for the institutions located in the same area I separated manually). Vertex size represents in-degree, vertex colour out-degree.

Table 4.2 shows a common pattern for females and blacks: their networks are sparser than those of white males. That is, they have fewer edges and so lower density, but also higher average path length, and a lower clustering coefficient. To a very great extent this is explained simply by the relative sizes of the three sub-populations.<sup>7</sup>

The geographic displays of the networks (Figures 4.1 and 18) show common hiring corridors, where universities with high number of connections in hiring tend to send their graduates to a high number of different institutions (nodes with high/low in-degree tend to have high/low out-degree); indeed, the correlation (excluding self-hiring) between in-degree and out-degree is 0.72 for SET and 0.53 in SSH.<sup>8</sup> I observe

apply throughout this text.

<sup>7</sup>The same patterns, and explanation, are present in SSH (Table 36).

<sup>8</sup>The exceptions to this strong correlation are for SET: Tshwane University of Technology, University of Fort Hare, and University of Limpopo that have a high connections in hiring (in-degree) and a low connections in placement (out-degree); while University of KwaZulu-Natal, and University of Cape Town instead show a high connections in placement (out-degree) and a low connections hiring (in-degree). See Table 34. For SSH I find the following exceptions: Walter-Sisulu University, Nelson Mandela Metropolitan University, and University of Limpopo with a high hiring connections (in-degree) and a low placement connections (out-degree); and University of Stellenbosch, and University of Cape Town with a high placement connections (out-degree) and a low hiring connections (in-degree).

### Summary Statistics SET hiring network

	All	Male	Female	White	Black
Number of Nodes	22	22	22	22	22
Number of Components	1	2	3	1	2
Number of isolated Nodes	0	1	2	0	1
Statistics on the Giant Component					
Number of Nodes	22	21	20	22	21
Number of Edges	133	115	57	107	52
Edge Density	0.288	0.274	0.15	0.232	0.124
Average Path Length	1.795	1.764	2.498	1.748	3.087
Diameter	9	6	12	9	15
Global Clustering Coefficient	0.648	0.588	0.511	0.578	0.41

**Table 4.2:** Summary Statistics SET hiring network for the years 1970-2004. Network statistics are computed without considering self-loops

that the universities with a high hiring and low placement connections are historically black universities, while those with an high placement and a low hiring connections are historical white universities, with the exception of KwaZulu-Natal.<sup>9</sup> This suggests that formerly black universities have been using hiring connections as a way to upgrade their faculty since the end of apartheid.

### 4.3.2 Prestige ranking

I consider university prestige as a social assessment, emerging from interactions among institutions. The well-known university rankings (THE, QS, Shanghai) can proxy prestige, but they remove interaction from the picture and so have questionable reliability, since prestige is not an individual attribute but it is part of a social process (Burris, 2004; Clauset et al., 2015). Consequently, following Clauset et al. (2015), I develop a measure of prestige ranking where institutional status arises from the patterns observed in the faculty hiring network. The algorithm starts from two hypotheses:

1. *Universities want to improve the quality of their research and teaching. A corollary is that they want to hire from universities that are “better” than themselves;*
2. *Scholars want to be hired by the best universities.*

<sup>9</sup>The University of KwaZulu-Natal was formed in 2004 by the merger of the University of Natal (white) and the University of Durban-Westville (Indian).

Were the desires expressed in these two hypothesis to be perfectly satisfied, and if the PhD institution is a reliable indicator of graduate “quality”, it would be possible to order universities (the rows and column of the adjacency matrix  $M$ ) so people only move down the ordering, implying that the adjacency matrix would have only zeros below the diagonal. Since actual hiring often departs from this “ideal” I search for an ordering that most closely approximates “zero weight below the diagonal”. To do this I apply to the adjacency matrix  $M$  an algorithm inspired by Vries (1998) and Clauset et al. (2015). The algorithm starts with a random ordering of rows (columns always having the same ordering as rows) of the matrix  $o_0$  and I compute the score  $s_o$  of this order, where:

**Definition 1.** An order  $o_k$  is an ordered n-tuple of universities names, and its score  $s_k$  is  $\sum_i \sum_{j>i} m_{ij}$ .

The algorithm tries to improve the score of the current order  $o_0$  using local search. Each iteration swaps two randomly selected nodes (both row and column). If the swap does not decrease the score I keep the swap, otherwise I reject it. After 100 iterations I stop, and record the resulting order and its score.<sup>10</sup> I repeat this procedure 10000 times to get a set  $O$  of 10000 orders and the set  $S$  of 10000 related scores.

**Definition 2.** The set of orders  $O$  is  $O = \{o_1, o_2, \dots, o_{10000}\}$ ; The set of scores  $S$  is  $S = \{s_1, s_2, \dots, s_{10000}\}$ .

Then I create the set  $Q$  of the orders  $o \in O$  with the maximum scores:

**Definition 3.** Let  $Q$  be the set of orders with maximum scores  $Q = \{o_k \in O | s_k = \max(s_k)\}$ , where  $O$  is a set of orders and  $k = 1, 2, \dots, 10000$ .

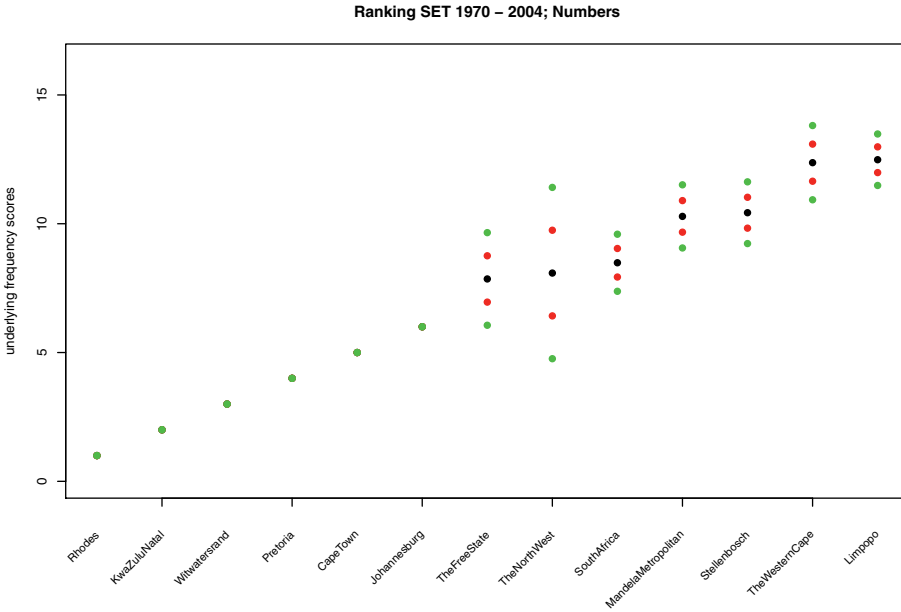
Then for each university I compute the mean of its ranks in the orderings in  $Q$ , to give the prestige scores, which provides a natural ordering or ranking of universities. Note that the prestige score is not a rank of natural numbers, it is an average value. This gives a better picture of university prestige where the distances in prestige among institutions are not of a fixed amount: pairs of universities adjacent in the ordering might have very different distances in terms of their scores.

I remove universities that graduated fewer than 5 PhDs in the period. Figures 4.2 and 19 show the results of the prestige rankings for SET and SSH. The frequency scores are in ascending order from the highest prestige (which corresponds to a score of one) to the lowest.

### Prestige ranking and other measures

To test whether the prestige hierarchy underlines by my measure is statistically different from a null model. I generate 5000 random matrices that preserve in- and out-degree, and also preserve the diagonal. This holds constant the number of PhDs each

<sup>10</sup>Checking manually, 100 iterations is almost always enough to find a local optimum.



**Figure 4.2:** Prestige Ranking for SET 1970-2004. The frequency scores are in ascending order: the highest ranked university has the lowest score. The black dots are the mean of the orders with the maximum scores in set  $Q$ , red and green dots are one and two standard deviation from the average. Universities with fewer than 5 PhDs are excluded.

university graduates, the number each university hires, and the number of graduates who are hired by their PhD institution. I apply the ranking algorithm to each of these 5000 matrices, and record the maximum score for each matrix. Taking these scores as the underlying distribution, I obtain a p-value of 0.002. This, I conclude that the hiring matrix has more structure than an equivalent random matrix, and that the hierarchy present in the empirical matrix is statistically stronger than those present in a random graph with the same characteristics. As a further robustness check I examine the correlations between prestige and measures of research output. From Web Of Science (WOS) I download aggregated data of research output for each South African university.<sup>11</sup> In SET the correlation of the prestige score with total publications 1988-2004 is 0.6 (p-value=0.04); with average citations per paper 0.5 (p-value=0.06); with total number of citing articles 0.5 (p-value=0.08); citations per capita 0.7 (p-value=0.01); and papers per capita 0.8 (p-value=0).<sup>12</sup> In SSH I find statistically significant correlations of prestige scores with average citations per paper 0.7 (p-value=0.01); with

<sup>11</sup>I select the university name in the address search, restricting to the period 1988-2004. For the universities that changed name I search pre and post merger names and I refine the WOS results looking at only journals publications. Last access February 2018.

<sup>12</sup>Per capita measures are obtained dividing total records by number of scholars in the database with a current affiliation in the particular university. This will over-estimate the per capita figures as there are (some, though few) South African authors in WOS who never apply for a rating. There is no reason to believe there is any bias in this over-estimate however.

SET PhDs hired from the top 5 universities

Rank	PhD University	Placed in SA academia	Placed in Top 5	Proportion placed in Top 5
1	Rhodes	38	31	0.816
2	KwaZuluNatal	115	85	0.739
3	Witwatersrand	113	102	0.903
4	Pretoria	165	120	0.727
5	Cape Town	158	116	0.734
	Total of top 5	589	454	0.771
Total PhD hires		1011		0.583

**Table 4.3:** PhDs hired from the top 5 prestigious universities in SET, according to the prestige ranking. The total of SET PhDs Hired in the period is 1011.

total number of citing articles 0.6 (p-value=0.05); and with citations per capita 0.6 (p-value=0.05). Thus this measure of prestige correlates well with measures of research performance, but is not identical to it.<sup>13</sup> This is in line with past contributions (Burris, 2004).

### Institutional Stratification

Tables 4.3 and 37 illustrate the institutional stratification hypothesis, for SET and SSH respectively. They show the number of PhDs from the top 5 prestige universities hired within the other top 5 institutions, and those hired in other institutions. The results are striking. In SET the top 5 prestige universities produce 58 percent of all PhDs within the country; among those, 77 percent find a first job within these 5 institutions.<sup>14</sup> This underlines the crucial role of prestige hierarchies in academia. Consistent with US-based work (Burris, 2004; Clauset et al., 2015), I find deep inequalities among universities in terms of first job placement: South Africa shows a pattern of stratification similar to those found in more mature knowledge systems. Moreover, the lower percentage of the first job placement of the top universities in SSH with respect to SET highlights the diverse hiring processes of the two fields. SSH are often governed by schools of thought and in South Africa language and culture also play important roles. This makes hiring processes more complex, more constrained, and less predictable.

<sup>13</sup>For SSH I find non statistically significant results for total number of publications and publications per capita. As discussed previously the methodology of this chapter is less suited for SSH.

<sup>14</sup>These numbers are 48% and 74% respectively for SSH.

## 4.4 Effects of rank change on future career

Changing university when taking up a first job after the PhD means moving within the prestige hierarchy. In particular, scholars move *Up* (*Down*) the hierarchy when hired by an institution with a higher (lower) prestige ranking than the one from which they received the PhD. And they *Stay* in the hierarchy when hired by a university of the same prestige of the PhD, which in practice means being hired by their own PhD institution.

Given the structural hierarchies in faculty hiring, people moving up the hierarchy are relatively rare, though it does happen. An obvious question to ask is whether this constitutes a signal regarding future career prospects. If the job market is able to identify promising (or weak) young scholars, a movement up (or down) would be predict a stronger (or weaker) future academic career. Further, it could also be that those who move up, relative to those who move down, on average work at higher prestige places, which could imply stronger colleagues and collaborators, and better resources. So there are *a priori* reasons to believe that movements in prestige between PhD and first job could be correlated with future research performance. But there lurks the issue of whether any link between prestige movements and performance is driven by individual quality or by the resources available at the receiving institution.

### 4.4.1 Matched pairs

To address this issue I do a matched pair analysis. I compare scholars' NRF ratings at different points in time (5, 10, 15 and 20 years after they were granted their PhD), asking whether people with different prestige transitions (Up, Down, or Stay) but having the same individual characteristics differ in rating. For each time span, I do a matched pair analysis comparing the transitions: Up vs. Stay, Down vs. Stay, and Up vs. Down. I match on gender, ethnic group, year of PhD, and either receiving or sending university. So, for each of the three comparisons I look at pairs of scholars from the same receiving or sending institution. When I match people with the same receiving university I compare people with same characteristics hired into the same institution, but having different PhD institutions. The match using the same sending institution, instead, compares individuals with a PhD granted by the same university, but who are hired in different places. To differentiate between sending and receiving institutions is important also because it is a control for possible Matthew effects on performance driven by university prestige. That is, the more prestigious a university is, the greater its ability to attract resources and this can result in higher productivity of the scholars and therefore higher NRF ratings. I solve this possible source of endogeneity by matching people with same receiving institution. Matching on the same sending institution controls for the fact that there is more scope for upward (downward) movement for those whose PhDs come from the bottom (top) of the prestige ordering. It also controls for training effects incurred during the PhD. The matching technique also controls for the other confounding factors on which I match: gender,

race and year of degree.

I use a re-sampling technique as follows. To do the Up vs. Stay comparison, I start with the set  $U$  of size  $N_u$  of people who move Up, and a set  $S$  of people who Stay. Then I sample with replacement,  $N_u$  people from the set  $U$  getting  $U'$ . For each member in  $U'$  I find the matches in  $S$ , and select one at random if there is more than one match. In this way I create matched pairs Up-Stay. Then I calculate and store the proportion of those pairs in which the Up person had a higher NRF rating ( $R$ ) than the Stay person:  $R_{up} > R_{stay}$ , and vice versa:  $R_{stay} > R_{up}$ . I repeat this procedure 10000 times, obtaining distributions of those proportions,  $F(p|R_{up} > R_{stay})$  and  $F(p|R_{stay} > R_{up})$ .

## 4.5 Results

The major concern of this chapter is whether prestige movements from PhD to first job correlate with the future research performance of the scholars. I compare, for different points in time, the NRF rating of individuals with the same characteristics but different prestige rank-change movements: Up vs. Stay, Down vs. Stay, and Up vs. Down. Since prestige could influence individual careers differently as time passes; I look at people's ratings 5, 10, 15 and 20 years after their PhD. For each comparison I study separately people hired by the same (receiving) university and those with a PhD granted by the same (sending) institutions. Moreover, the matching technique accounts also for additional confounding factors: gender, ethnic group, and PhD obtained year. As I am interested in the effects of rank change, the null hypothesis, **H0**, is that prestige movements are unrelated to research performance:

$$\mathbf{H0}: F(p|R_{up(or\ down)} > R_{stay}) = F(p|R_{stay} > R_{up(or\ down)}). \quad (4.1)$$

Performing two-sided KS tests, I find that this is not the case; in each comparison the distributions are different. Given that the distributions are different from each other, I can ask whether one stochastically dominates the other (less and greater one-sided KS tests). The results for the tests are in Appendix 38.<sup>15</sup>

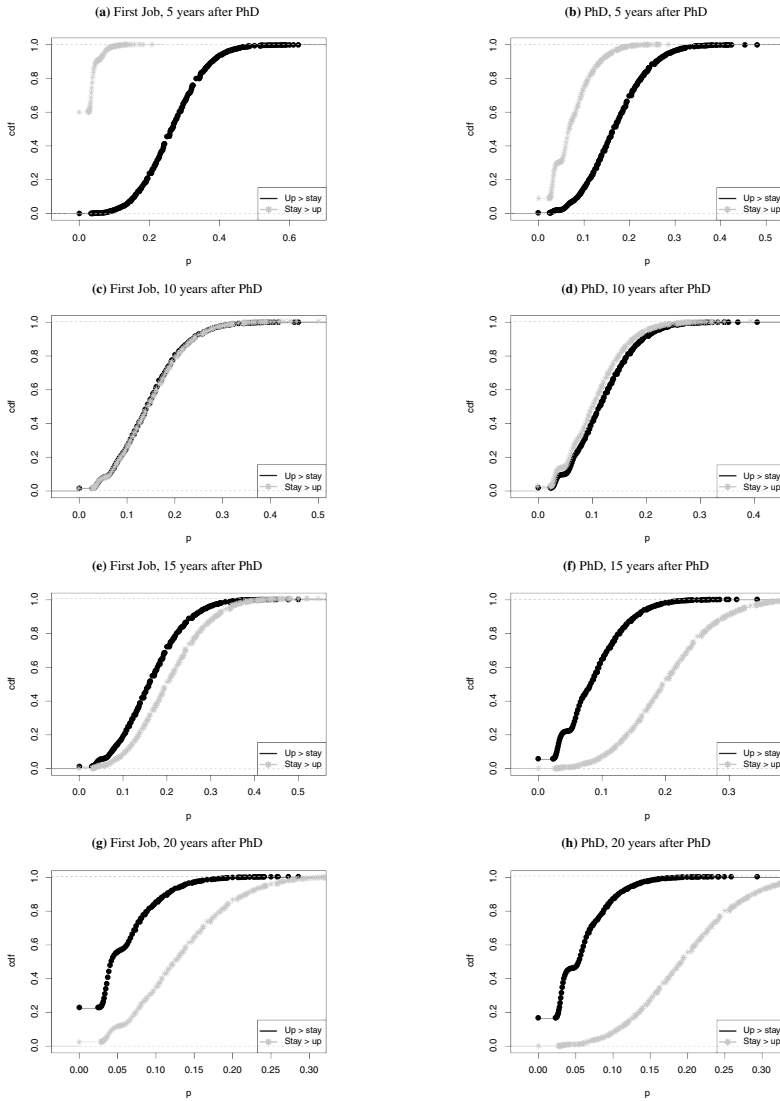
The definition of first order stochastic dominance is typically given as

**Definition 4.** a CDF  $F(x)$  First-Order Stochastic Dominates  $G(x)$  iff  $F(x) \leq G(x)$  for all  $x$  (Mas-Colell et al., 1995).

Figure 4.3 shows the results of Up vs. Stay. Looking at people hired by the same institutions but different PhDs (left column), and at those with the same PhD institution but different first jobs (right column), I find consistent results. After 5 years  $F(p|R_{up} > R_{stay})$  (black curve) stochastically dominates the other  $F(p|R_{stay} > R_{up})$  (grey curve); while 15 and 20 years later I have the reverse:  $F(p|R_{stay} > R_{up})$  dominates  $F(p|R_{up} > R_{stay})$ . This tells a consistent story. In the short term, those

<sup>15</sup>Additionally I perform a robustness check following a bootstrap technique in appendix C.21.

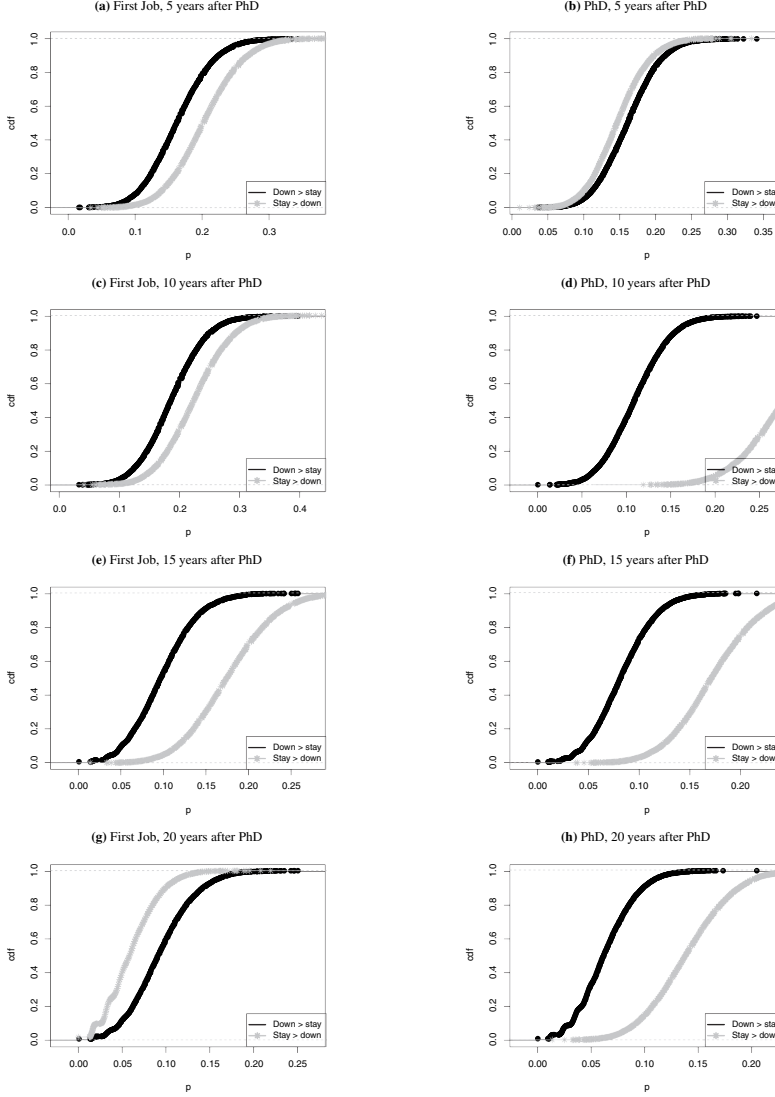




**Figure 4.3:** Up versus stay comparison. The black curves are cumulative distribution functions of the proportion of observations in which  $R_{up} > R_{stay}$  was the case for  $p\%$  of the matched pairs. Grey curves are the CDFs for the  $R_{stay} > R_{up}$  proportions. From top to bottom 5,10,15, and 20 years after PhD. Pairs matched using gender, race, PhD obtained years and first job university (left column) or PhD institution (right column).

who are promoted, moving up in prestige, have higher ratings than do those who stay; but in the long term the opposite is true: those who do not move (stay) have higher ratings. So in the long term, looking at people with the same first job, I find that those with “better” PhDs (stay) do better; while looking at people with same PhD institution, those with “worse” jobs (stay) perform better. Intuitively, the former seems reasonable, the latter odd. I return to this below.

Figure 4.4 shows the results for the Down vs. Stay comparison. With two excep-



**Figure 4.4:** Down versus stay comparison. The black curves are cumulative distribution functions of the proportion of observations in which  $R_{down} > R_{stay}$  was the case for  $p\%$  of the matched pairs. Grey curves are the CDFs for the  $R_{stay} > R_{down}$  proportions. From top to bottom 5, 10, 15, and 20 years after PhD. Pairs matched using gender, race, PhD obtained years and first job university (left column) or PhD institution (right column).

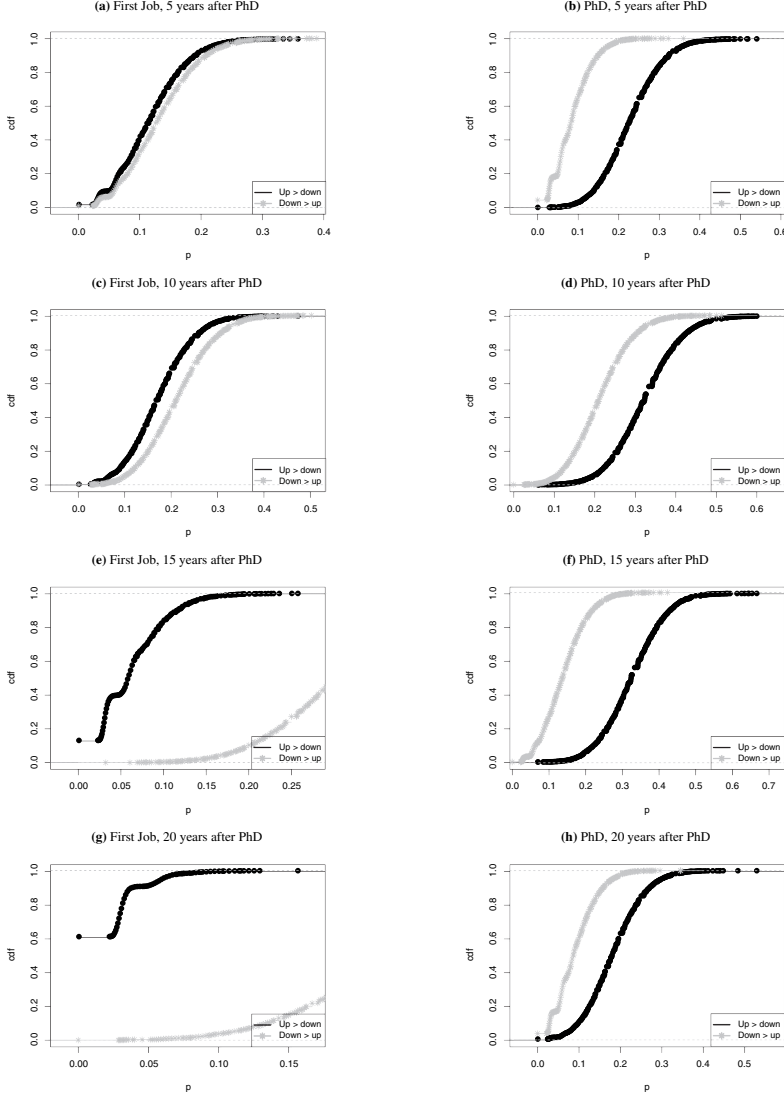
tions (first job match 20 years, and PhD match 5 years) I find that  $F(p|R_{stay} > R_{down})$  stochastically dominates  $F(p|R_{down} > R_{stay})$ . That is: those who stay have higher ratings than those who move down in prestige. In particular looking at those with the same first job (left column) I have that those with “worse” PhDs (stay) do better; while when I look at people with same PhD institution (right column), those with “better” jobs (stay) perform better. Again I find that in general those who do not move after their PhD have higher ratings.

In the case of Down versus Stay, matched on PhD, the value of staying seems reasonable: of two people with equivalent PhDs, the one who stays will have the higher prestige job and a higher rating. Similarly in the case of Up versus Stay matched on first job: of two people with the same job, the one with the more prestigious PhD (the one who stays) will have a higher rating. However, Down-Stay matched on job, and Up-Stay matched on PhD seem somewhat paradoxical, as in the first case the less prestigious PhD (with the same job) does better, and in the second the less prestigious job (with the same PhD) does better. I discuss this apparent paradox below.

Figure 4.5 shows the Up versus Down comparison. It should be noted that because in both groups the sample is relatively small, this comparison is less reliable, and needs cautious interpretation. To have a reasonable number of matches I relax the matching of PhD obtained year, here considering an interval of six years. That is, two agents match on PhD year if they are within  $\pm 3$  years of each other. Figure 4.5 shows, looking at researchers hired by the same institution (left column), that  $F(p|R_{down} > R_{up})$  dominates  $F(p|R_{up} > R_{down})$ . Those moving down to a job from a higher prestige PhD institution do better than those moving up to the same job from a lower prestige institution, and the gap increases over time (the distance between the two cumulative distributions increases). This seems reasonable: all else equal, a higher prestige PhD is good. Matching people with the same PhD institution (right column) I find that  $F(p|R_{up} > R_{down})$  dominates  $F(p|R_{down} > R_{up})$ . That is, those who move up in prestige (hired by a more prestigious institution) perform better in ratings than those hired by less prestigious universities. This also seems reasonable: all else equal, a higher prestige job is good. So comparing people who experience mobility in the transition from PhD to first job is found that, holding job constant, coming from “better” PhD (down) is good; while holding PhD constant, going to “better” job (up) is good.

To summarise the results in a more intuitive way: when I compare people who experience transitions from the PhD to first job with those who do not (that stay) I find a beneficial inertia effect. Generally speaking, those who stay in the same university after the PhD have higher rating. This positive effect of inertia might be due to various factors. In some instances training effects from the PhD work to support the result: those trained at a particular institution will have skills and expertise which fit well into the research taking place there, and perhaps complement that of the institution very well. This argument may be particularly pertinent in a small country like South Africa, where not all universities are expert in all fields. There is considerable specialization and “division of labour” among the universities regarding research fields, implying that in many disciplines there are only few institutions with strong research presence. This would also explain the heavy diagonal in the hiring matrix.

However, the value of institutional inertia may also simply relate to the way the PhD job market works. A university has better information about its own graduates, and so it can make better judgements with respect to their intrinsic quality. Moreover, any research-focussed university will have strong incentives to keep its best graduates



**Figure 4.5:** Up versus down comparison. The black curves are cumulative distribution functions of the proportion of observations in which  $R_{up} > R_{down}$  was the case for  $p\%$  of the matched pairs. Grey curves are the CDFs for the  $R_{down} > R_{up}$  proportions. From top to bottom 5, 10, 15, and 20 years after PhD. Pairs matched using gender, race, PhD obtained years and first job university (left column) or PhD institution (right column).

so as to enhance its own reputation rather than that of its rivals. The comparison of Up vs. Down instead tells the role of university prestige; when people experience PhD to first job mobility their movements in prestige are crucial: holding PhD constant, moving to a more prestigious first job is better; and holding first job constant, having a more prestigious PhD is better. It appears that training and resources (broadly defined) both contribute.

### 4.5.1 Possible cohort effects

The results show that there are some differences in the effects of prestige movements depending whether the scientist is observed 5, 10, 15 or 20 years after PhD. One possible explanation is cohort effects. These could drive this observation because the cohort composition is different in each time sample. Specifically, since the rating data run only from 1983 to 2012, observations on the ‘5-year’ ratings include PhDs from 1970 to 2004; whereas 20-year ratings include only PhDs from 1970 to 1992. To test for this cohort effect I repeat the analysis restricting the sample to only older scholars, with a PhD granted before 1992. Figures 15, 16, and 17 in Appendix C.18 show the results for the restricted sample. These results are consistent with the ones discuss above, so I can exclude this hypothesis: the results are not driven by changes in the cohort composition.

### 4.5.2 Size effects

PhD production and graduate hiring are very skewed, and some universities are bigger than others. The prestige ranking looks not only at faculty production but it considers in particular the quality of placement. So a small institution could achieve high prestige by placing a small number of graduates at high prestige universities. In order to exclude the possibility that size effects drive the measure I re-run the algorithm using the logs of the entries in the matrix. This is a monotonic transformation of university sizes but makes the differences between them much smaller, so results are less likely to be driven by size alone. The prestige ranking results are almost identical. Further, the correlation of out and in-strength of the adjacency matrix, is not statistically distinguishable from zero ( $\text{cor}=0.3$ ;  $\text{p-value}=0.3$ )<sup>16</sup>. This implies that amounts of hiring and placement are unrelated, so universities with high production of PhDs are not the ones that hire more. One other suggestion that the results are not driven by size comes from the fact that Rhodes University, consistently ranked first in prestige, is much smaller than the other “top 6”.

## 4.6 Discussion and Conclusion

This chapter reveals important aspects of how the PhD job market works in South Africa and the results tend to be in line with previous work. As is often observed in other settings, I find that institutional stratification in higher education holds in the South African context. The 5 most prestigious universities produce between 48 and 58% of all PhD graduates (who enter academia) in the country and they tend to hire graduates from this elite group. Occupational segregation is also present in South Africa as elsewhere: under-represented groups are less likely to get jobs in higher prestige universities than are white males.

---

<sup>16</sup>This is done removing the diagonal of the matrix

Looking at the relation between prestige transitions from PhD to first job and individual research performance I find two main results. On the one hand, comparing people who experience a prestige transition with those who do not, I find a positive inertia effect. Those who stay in the same institution after the PhD have higher performance than those who move. At first glance, what appears to matter is not moving up or down the prestige hierarchy, but rather resting in an established environment. However, when I compare explicitly those who make upward and downward transitions, I find that university prestige is deeply related with academic performance, consistent with previous literature (Burris, 2004; Clauset et al., 2015). Holding PhD prestige constant, those with a more prestigious first job have better long run performance than those with a less prestigious first job. But similarly, holding first job constant, those having a PhD from a more prestigious institution have better long run performance than those with a less prestigious PhD. This suggests that prestige is at the very least a signal of quality, but possibly also has causal effects. More prestigious PhD programmes may attract better students and or give better training.

The results in this chapter underline the big role played by inertia in the South African PhD job market. In the data, of the rated researchers in South Africa, historically roughly two thirds of those going into the professoriate do not change institutions at the completion of the PhD.<sup>17</sup> This is maybe related to culture, institutional organisation, and history. But it seems likely that it will disappear over time. In many locations it has been common in the past, but today in North America (or the anglo-saxon world more generally) it is rare that a department will hire its own graduates, and while still more common in Europe, it appears to be disappearing, even in countries like France and Italy.

I observe that those who do not change institutions after the PhD tend to have higher NRF ratings later in their career. To return to the discussion in the introduction, this suggests that imperfect information is fairly severe in PhD hiring: universities have good information about their own graduates, and can successfully “pick the winners”. Of their own students, they have better information on candidates’ intrinsic qualities than do other institutions, and so can make better judgements. Further, they have strong incentives to encourage their best students to stay, rather than to let them drift away to strengthen competing institutions. This is consistent Robert Merton’s observation that “*leading universities manage to identify early, and to retain to their faculties, those scientists of exceptional talent: they keep 70 percent of future [Nobel] laureates, in comparison with 28 percent of other Ph.D.’s they have trained.*” (Merton (1968) page 7).

To understand why those internally selected perform better over a long period of time this chapter suggests to look further at the deep causes of the positive role played by inertia. The role of inertia might be related not only to the general dynamics of the

<sup>17</sup>In spite of the 2004 reforms, this feature remains part of the SA academic world: in the period 2004 to 2012 it still the case that about two thirds of graduate receive their first appointment at their PhD institution. The fraction is slightly higher in SET, slightly lower in SSH.

university system and the PhD job market, but might also be linked to the behaviour of the scholars in terms of co-authorship and specialization. In particular, young researchers who do not experience mobility may have different collaboration patterns than do their counterparts. They can have more stable co-authorship linkages able to sustain their careers, especially in the early stages. Coupled with the Matthew effect this would create both short and longer term positive effects of inertia. Furthermore, the research orientations of those who stay and those who move may be different in terms of specialization. Those internally selected could be more specialized in a particular area of research that is more germane to the home department, and this specialization could drive their long-run performance. This would be true in any small country whose institutions have specialized in particular areas. This type of further research at micro level could shed new light on the university system not only of South Africa but also of other countries with low first job mobility.

It is important to realize that this is an historical analysis. The most recent PhD in the data is from 2004, and the most recent ranking from 2012. Indeed, one limitation of the chapter is that consider the hiring network and consequentially the prestige hierarchy among universities as fixed across time. In contrast, the South African university system experienced large changes as for example the reform in 2004. The research performance of universities changed substantially over the period. In particular, among the formerly black universities there have been several notable changes in research output (University of Fort Hare or the University of the Western Cape for example), suggesting that some of these universities may be entering a different era and playing a different role in the system. However, there is little reason to believe that information asymmetries surrounding PhD hiring will disappear any time soon, in South Africa or elsewhere, so the results regarding effects of mobility on career success are likely to be robust.

Because in principle universities have as their *raison d'être* the creation and diffusion of knowledge, and because by its nature knowledge changes relatively slowly (that is to say, what is true does not change quickly) universities tend to evolve relatively slowly, and so do their standings relative to each other.<sup>18</sup>

Seen from this perspective, stratification, and to a lesser extent inertia, in hiring is a natural outcome. These two forces are a source of the Matthew effect at the university level, and tend to create stasis, or possibly even reinforce the gaps in university hierarchies. Whether or not in general a hierarchical or even two-tier university system is good or bad, in the South African context where the current hierarchy is born of the apartheid period, one can argue that the existing hierarchy is not ideal. The top universities in the current structure tend to be the historically white universities, and there, even 20 years after apartheid ended, the professoriate remains predominantly white, and for structural reasons is likely to remain so for many years (Barnard et al 2017). Here the hierarchy of the universities is socially problematic. In this respect

---

<sup>18</sup>For a discussion of this see for example Cowan et al. (2010) pp. 278-299

a policy devoted to increasing the in-house capabilities of the latter and the exchange of expertise between universities could help to reverse this trend, and to create a more equal and productive system.





## Chapter 5

# University prestige in the Ph.D. job market and the big-fish-little-pond effect

This chapter is a natural expansion of the previous one. Here I investigate how the hierarchical stratification of university prestige affects scholars' career achievements focusing on 898 STEM Ph.D. graduates that enter into their first academic job in Mexico between 1992 and 2016. The previous chapter shows how prestige and mobility are important aspects of academic life and how inertia plays a critical role during early-career. Indeed, job market candidates compete for positions at universities but many of them have thin research records, making signals of quality such as the prestige of the Ph.D. institution of crucial importance. I further develop the technique described in the previous chapter to rank university prestige based on the hiring networks among universities. In order to overcome a possible drawback of the methodology, here I use a dynamic version of the ranking algorithm, which includes new information in the job market and prestige dynamics. Finally, I use the quasi-experimental method based on matched pairs of individuals developed in the previous chapter. The chapter confirms the role of inertia also in the Mexican context — scholars hired internally sustain higher performance over their careers in comparison to those who move the hierarchy. Interestingly, the chapter shows that scholars who experience upward prestige mobility exhibit, on average, lower academic performance than the other groups. I will argue that the negative (positive) relation between upward (downward) prestige mobility and performance can relate to a “big-fish-little-pond” effect (BFLPE) or observationally equivalent mechanisms that underline the need for further research. The evidence of a BFLPE-like effect has policy implications because it can hinder the flows of knowledge throughout the science system and individual achievements as well.

---

<sup>0</sup>This chapter is based on Gonzalez-Sauri and Rossello (2019), UNU-MERIT working paper series.

## 5.1 Introduction

Training, collaboration, and mutual evaluation characterise an academic profession. Those tasks imply that scholars know each-other and/or their work and contribution. In those circumstances, networks among scholars and institutions naturally emerge. However, a prestigious position in such networks is more likely to generate positive network effects. In particular, Matthew effects can enhance individual performance, productivity, and prominence (Merton, 1968; Bol et al., 2018; Teplitskiy et al., 2020; Horta et al., 2018). To decompose performance and endogenous network effects, sociology, economics, and science of science examined the relationship between the prestige of the university affiliation of an individual and her/his performance (Fox, 1983; Headworth and Freese, 2016; Su, 2011). Those studies highlight that university prestige is positively associated with individual performance. But, at the same time, results appear mixed when the endogeneity of university prestige is accounted for (Allison and Long, 1987; Williamson and Cable, 2003; Bair, 2003; Miller et al., 2005; Laurance et al., 2013; Lawson and Shibayama, 2015). In particular, much of the evidence accumulated so far indicates that prestige from the Ph.D. institution increases the chance of getting further prestigious appointments. Under certain circumstances, Bedeian and Feild (1980) even suggest that institutional prestige from the Ph.D. might be more important than research records before the first-job appointment. However, these results are to a certain extent expected, science systems exhibit Matthew effects, where a “good” initial affiliation is likely to be followed by other prestigious appointments and higher academic performance.

These studies have been conducted in top-tier institutions, mostly in North America (Clauset et al., 2015). Those universities often have a long history and are well known and integrated internationally (Demeter and Toth, 2020). In contrast, less mature knowledge systems have higher resource constraints and many institutions are relatively new and less known. Thus, the presence of a star scholar as well as winning (or not) a grant may change prestige and output of an entire department (Azoulay et al., 2012). In similar circumstances, prestige can be more volatile and the distance between prestige and “quality” could be potentially larger compared to mature university systems.

The gap between prestige and “quality” can also change for individuals throughout their careers. Such a gap is potentially larger when signals of quality matter the most. After Ph.D. graduation early-career scientists have often thin publication records (Pifer and Baker, 2013). Additionally, a senior faculty member and her/his lab often supervises the student’s doctoral research. Thus, hiring PhD job market candidates is potentially problematic. On the one hand, information about abilities is noisy making signals of “quality” relevant for hiring decisions (Cowan and Rossello, 2018). On the other hand, the sorting of candidates done by the Ph.D. job market could have long-run consequences because of Matthew effects. In this respect, past

research highlights that Matthew effects and cumulative advantages start early in the career (Bol et al., 2018; Lee, 2019). Therefore, the transition into the first job and the following years might have consequences for scholars future careers (Bazeley, 2003).

At the same time, for universities, forming and hiring Ph.D. graduates is essential for their competitiveness — young scholars will fill academic positions over a long period. Indeed, the Ph.D. job market involves the decisions of both Ph.D. graduates and hiring committees at universities.

All these decisions are embedded in the university hiring networks of Ph.Ds. The university hiring network has been extensively studied and generates predictable hiring flows from one university to other (Barnett et al., 2010; Clauset et al., 2015; Cowan and Rossello, 2018; Lang et al., 2019). Indeed, past research shows that the university hiring networks tend to contain an implicit hierarchical order of prestige among institutions (Barnett et al., 2010; Mai et al., 2015). This hierarchy of prestige comes to existence because hiring decisions in the academic labour market are pairwise evaluations of “quality” between candidates and universities. I take advantage of this pairwise assessment of quality and I use it to extract information about the distribution of prestige among universities (Cowan and Rossello, 2018; Clauset et al., 2015). However, often such methods do consider prestige as fixed across time. In contrast, prestige may change according to modes, tastes and new information. If for example, frauds and misbehaviour are found within a prestigious department its reputation will decrease. In the same way, if a department wins a large National grant its reputation will improve. In contrast with past research, this chapter proposes a ranking algorithm that updates the new information contained in the university hiring networks. Thus, I make dynamic the ranking algorithm proposed by (Cowan and Rossello, 2018). In this way I account for various confounding factors driven by the possibility of new information contained in the Ph.D. job market.

This chapter asks how prestige stratification affects labour market mobility and future academic performance in Mexico. Most of past literature assumes that prestige is static and that has on performance an absolute effect. In other words, individuals perform at their best in prestigious environment irrespective of their past experiences. However, the effect of prestige on performance can be relative for the individual to the movements she/he does in the hierarchy. This pattern seems to be more complex and nuanced. The previous chapter, studying prestige stratification in South Africa, finds that scholars who experienced upward prestige mobility exhibit lower academic performance in comparison to their Ph.D. fellows who kept their position in the hierarchy.

Indeed, after graduation, scholars can move to other institutions or be hired internally. When hired internally, researchers keep their current social capital and hold their position in the prestige hierarchy. In contrast, when they move, they experience upward (downward) prestige mobility — they are hired in a university more (less) prestigious compared to their Ph.D. Analysing these prestige changes rather than solely the

effect of the Ph.D. granting institution can offer a deeper insight into how university prestige stratification relates to early-career and academic performance.

Those prestige changes may involve numerous psychological mechanisms and the working environment experienced by agents. Those mechanisms relate to how individuals perceive their place in this prestige hierarchy and relate to one other. Indeed, beyond the absolute link between prestige and performance, we should consider that prestige can affect individual performance also in a relative way. In particular, the effect that prestige has on individual performance may relate to the “profile of prestige” an individual is accustomed to.

For example, the psychological literature finds the big-fish-little-pond effect (BFLPE), where individual performance relates to the average performance of her/his peers. That is agents might perform “better” ( “worse”) in an environment with a “lower” (“higher”) competition where the perceived average ability of peers is lower (higher). Testing these psychological mechanisms is beyond the scope of this work, but I highlight a non-linear association between prestige and performance (see section 5.8 for more details).

With these premises, this chapter incorporates a dynamical representation of prestige across time and the idea that university prestige and its effect on performance is relative for each individual. In particular, I test the effect that the movements in the prestige hierarchy have on individual performance. Based on a sample of 898 Mexican Ph.D. graduates in STEM between 1992 and 2016, I find a “lock-in” where Ph.D. candidates hired internally exhibit higher academic performance throughout their careers than those who move up or down the hierarchy. Further, concentrating the analysis on large movements in the hierarchy I find evidence for a sort of BFLPE where an upward (downward) prestige mobility in the transition into the first academic job negatively (positively) relates with academic performance. Both results underline a complex association between prestige stratification and individual performance

## 5.2 Prestige, mobility and performance

In section 4.2 the previous chapter reviews the literature on university prestige. Here I provide a brief summary

### 5.2.1 University prestige

The main approaches for ranking universities by prestige are input-output (Debackere and Rappa, 1995; Chan et al., 2002; Kalaitzidakis et al., 2003; Oyer, 2008; Buéla-Casal et al., 2012), survey (Abbott and Barlow, 1972; Cyrenne and Grant, 2009; Moodie, 2009; Olcay and Bulu, 2017), and network-based measures (Barnett et al., 2010; Cowan and Rossello, 2018; Zhu and Yan, 2017; Nevin, 2019).

Network-based measures are a new framework to rank universities that examine a social process in which scholars recognize quality in their work and this evaluation is cross-validated by the interactions between institutions. Following these lines, the ranking approach in this chapter builds upon the idea that the Ph.D. hiring networks contain information about how agents evaluate each-other quality (Clauset et al., 2015). I apply the new methodology proposed by Cowan and Rossello (2018) to study the effect of prestige stratification on academic performance in the context of the South African science system. They consider university hiring patterns as revealing of pairwise assessments of quality between Ph.D. graduates and hiring committees. In other words, since hiring implies underline decisions for agents, they exploit the information contained in the movements in the labour market to approximate the distribution of prestige. As a further expansion of their methodology, this ranking algorithm is dynamic across time. Thus, I consider the job market of variable sizes that consequentially changes the prestige of the institutions involved. A dynamic perspective to rank universities is more suited in a large developing setting where the distribution of prestige is potentially more volatile.

### 5.2.2 The link between mobility and performance

The literature studies the relationship between university prestige and mobility to understand how the prestige of the Ph.D. granting institution affects the placement and subsequent labour market outcomes. In the Ph.D. job market, both individuals and universities have incentives to make accurate choices.

However, the asymmetry of information in the Ph.D. job market makes the prestige of the Ph.D. granting institution a signal for the unobserved skills and abilities of applicants. Indeed, past literature shows a positive relationship between the prestige of the Ph.D. granting institution and future employment (Crane, 1970; Debackere and Rappa, 1995; Bedeian et al., 2010; Lawson and Shibayama, 2015; Pinheiro et al., 2017; Headworth and Freese, 2016). Ph.D. graduates from prestigious institutions tend to get “better” job compared to graduates from lower-tier universities. Similar results suggest that a prestigious Ph.D. function as a key mechanism to alleviate the asymmetry of information between candidates and hiring faculties. Moreover, since a “good” affiliation provides opportunities, networks, and resources this first sorting of the job market has potential long-run consequences on career achievements (Oyer, 2008; Bedeian et al., 2010). A similar early-career advantage is potentially problematic when quality and prestige do not entirely overlap. In this respect, some studies argue that institutional prestige from the Ph.D. granting institution is more important than researchers “quality” for obtaining a first academic job (Long et al., 1979; Allison and Long, 1990; Baldi, 1995; Gerhards et al., 2018).

Another group of studies pay attention to how changes of institutional prestige relate to academic performance and academic achievement of scholars. Oyer (2008) use a longitudinal sample to estimate how changes in institutional prestige affect the

academic performance of economists. He shows that even after controlling for proxies of individual-level ability, that early academic prestige positively correlates with academic performance measured by publication productivity. Moreover, he shows that scholars generally move down the prestige ranking over their careers. He argues that this is because high ranked universities produce a significant percentage of the total graduates that later move to lower-ranked universities. Chan et al. (2002) examine the mobility of scholars publishing in 16 top financial journals. They find that upward ranking mobility is rare and that scholars who experienced it produce twice as many publications compared with average production of scholars from destination universities. He further shows that after controlling for ability using publication productivity, the rank of the Ph.D. granting institution predicts upward ranking mobility through their academic careers. Azoulay et al. (2014) take an alternative approach, comparing academic performance of scholars before and after upward mobility given by a prestigious academic recognition. They find that gains from upward ranking mobility have a lower effect on scholars that have above average citations than on scholars with low or below average citations. In general, these studies suggest that upward ranking mobility is associated with higher academic performance, but this is not always the case. Cowan and Rossello (2018) offer a closer look at how prestige differentials from the Ph.D. to the first job relate to academic performance. They use a quasi-experimental methodology based on matching pairs examining a sample of 1011 South African job market candidates in STEM. They show that scholars hired internally (maintaining their place in the prestige hierarchy) exhibit on average higher performance compared with scholars that move up in the prestige rank. Their work underlines that the link between prestige changes and performance might be complex and non linear. In the following subsection I discuss psychological mechanisms that can play a role.

## 5.3 Data

Data originates from the Mexican National Council of Science and Technology (CONACYT). Data were collected through the most extensive science policy of the country, the *National System of Researchers* (NSR), whose aim is to increase the productivity, quality and competitiveness of Mexican researchers. NSR was implemented in 1985 when the primary motivation behind the policy was the rising concern about technological capabilities and performance of the Mexican science system under the threat of inflation and budget cuts. Reyes Ruiz and Suriñach (2015) describe how the policy evolved across the years, but in general, its structure is substantially unchanged.

The general framework of NSR is the following. Each researcher applies to the NSR submitting her/his curriculum vitae and publications and is assigned to one of seven different research disciplines. For each discipline, the NSR forms a committee of prominent researchers that evaluates the performance of applicants following a peer-review process (described in Appendix D.24). In contrast to bibliometric mea-

tures, a peer-review evaluation has the advantage of including a holistic evaluation taking into account the validation practices of each discipline within the country. The evaluation process ends with a rating that systematises the academic performance of researchers in 5 ordered categories. In the chapter, I use those categories to measure the academic performance of individuals.

I focus on STEM graduates excluding from the analysis those in Social Sciences and Humanities to reduce the potential influence that schools of thought have on the Ph.D. job-market.<sup>1</sup> The sample spans 25 years representing 898 Ph.D. job-market candidates hired in a Mexican university between 1992 and 2016. I include 36 Mexican institutions<sup>2</sup> and *longitudinal* records for each scholar of academic performance and individual level controls as *gender*, *discipline*, *graduation year* and *evaluation year*.

## 5.4 Interactive Prestige Ranking

In this section, I describe the interactive algorithm based on universities hiring networks that I use to rank Mexican universities according to their prestige. This method is based on the idea that movements in the academic job-market contain information about how universities and Ph.D. graduates perceive each other's quality.

The hiring network,  $G = (V, E)$ , represents the movements in the Ph.D. job-market. The vertices  $V$  of the network are the universities participating in the job-market and edges  $E$  represent movements of Ph.D. graduates from one university to another.  $A$ , in Table 5.2, is the weighted directed adjacency matrix of the hiring network representing the flows of graduates from Ph.D. to their first-job institution. Each entry in the matrix ( $a_{ij}$ ) represents number of Ph.Ds. where the diagonal elements ( $a_{ii}$ ) are the Ph.D. graduates hired internally (trained and hired by their faculty) while the off-diagonal entries ( $a_{ij}$  with  $i \neq j$ ) are graduates from university  $i$  hired as their first academic job in university  $j$ .

Following the previous chapter, the ranking algorithm is based on two key assumptions. The first concerns universities, that is, they try to improve their status and quality and in pursuing it they try to hire from universities “better” than themselves. The second assumption considers scholars, and they want to be hired by the most prominent institution. When both academics and institutions satisfy those desires fully, in  $A$  exists a unique order of university names (rows/columns names) such that Ph.D. graduates only move down the hierarchy. In other words, under this assumption, rows and columns of the adjacency matrix  $A$  can be rearranged in an *upper triangular matrix* such that all entries below the diagonal are equal to zero. I define this unique order  $o^*$ , where the sum of rows have a global maximum score equal to  $s^*$ .

<sup>1</sup>CONACYT Research Areas included are I, II, III, VI and VII. Respectively, Physics-Mathematics and Earth Sciences, Biology Chemistry and Life Sciences, Medicine and Health Sciences, Biotechnology and Agricultural Sciences, and Engineering. See Appendix D.24.

<sup>2</sup>The list of institutions is presented in the Appendix D.22



$$s^* = \max \sum_i \sum_{j>i} A_{ij} \quad (5.1)$$

Geographic location and other recruiting criteria imply that the Ph.D. job-market often departs from this strict assumptions. Thus empirically the order is not unique and  $A$  is not a perfect *upper triangular matrix*. However, since prestige is an important selection criteria for university and scholars I apply the heuristic algorithm proposed in the previous chapter to find the set of orders that gets as closer as possible to  $s^*$  and have the minimum number of violations from a upper triangular matrix configuration.

In particular, as described in the previous chapter, to approach the underlying prestige hierarchy the algorithm works at follows. The algorithm for  $k = 10000$  times starts assigning to  $A$  a random order of rows then it computes the score  $s_k$  such that

$$s_k = \sum_i \sum_{j>i} A_{ij} \quad (5.2)$$

For 100 times the algorithm tries to improve the score  $s_k$  in the following way. For each iteration two nodes (both rows and columns) are randomly selected and swapped. If the swap does not decrease the score it keeps it, otherwise it rejects it. After this 100 searches the obtained order  $o_k$  and score  $s_k$  are recorded, obtaining a set of  $n$ -tuple  $O = \{o_1, o_2, \dots, o_k\}$  orders and their associated scores  $S = \{s_1, s_2, \dots, s_k\}$ .

From these two sets  $O$  and  $S$ , the algorithm selects the set  $Q = \{o_m \in O \mid s_m \approx s^*\}$  of orders that reach the highest score. Each  $n$ -tuple of the set  $Q = \{q_1, \dots, q_m\}$  contains a possible university rank  $R(v) = \{r_1, r_2, \dots, r_m\}$ . Then for each university the algorithm computes its prestige score according to the formula

$$R_v = m^{-1} \sum_{i=1}^m r_i \quad (5.3)$$

which is in other words the mean of its ranks in the set of orders with maximum score  $Q$ .<sup>3</sup> The prestige score of each university provides a natural ordering or ranking of universities that is the measure of prestige.

A key assumption of this algorithm is that a single adjacency matrix  $A$  captures the underlying hierarchy of prestige. This implies that in all universities (nodes) and scholars participate in the labour market and the size of the market is fixed. However, movements between universities over an interval of time can be constrained by various forces.

The assumption of a fixed hiring network to measure prestige is consistent with previous evidence that finds prestige to be persistent across time. However, in a less developed higher education system prestige might be more volatile. Trends, winning a grant or a famous scholar can highly influence the prestige of such institutions. To account for these possibilities, I relax the assumption that the hiring network is fixed

---

<sup>3</sup>Further details of the procedure are described the Algorithm 1 from Appendix D.23.

across time adopting a *dynamic* computation of the algorithm. The proposed variation iterates the previous algorithm over closed intervals,  $t = [y - \Delta, y + \Delta]$ , of time centred around the Ph.D. graduation year  $y$ , with fixed windows of  $\Delta = 3$  years. This implies that the hiring network and the scholars and universities involved are different for each time window.<sup>4</sup> However, not all institutions are present across  $t$  intervals, for instance, more recent universities are not listed in the early years of the sample. Hence, the final scores of the *Interactive Prestige Ranking* is the average score of each university  $i$  over  $t$  intervals of time.

After the computation of the ranking, I distinguish between three groups of scholars, *Up*, *Down* and *Stay*, that exhibited different changes in the prestige during the transition between Ph.D. graduation and first job in the following way. For each researcher in the sample, I calculate the difference between the Ph.D. prestige rank and the one of her/his first job institution. The difference is positive (negative) for the group of scholars that move *Up* (*Down*) which experience upward (downward) prestige mobility — they are hired by a university more (less) prestigious than their Ph.D. The difference is equal to zero for scholars who *Stay* experiencing internal hiring — those hired by their Ph.D. institution.

## 5.5 Prestige Ranking Results

Table 5.1 shows the ranking of Mexican universities using the dynamic ranking with a 3 years time window.

The aggregate adjacency matrix in Table 5.2 illustrates the mobility from the Ph.D. to first job for the complete sample over the entire period. The matrix displays a heavy diagonal representing Ph.Ds. hired internally. Looking at the movements in the prestige hierarchy, Table 5.3 highlights a high level of stratification in the Mexican university system. Where the 10 most prestigious Mexican universities produce the 68% of Ph.D. graduates and nearly half of them are hired as a first-job in those institutions.

---

<sup>4</sup>The number of universities is very small in early years of the sample their analogous adjacency matrices yield trivial orders of small length. To overcome this, I aggregated the first years of the sample up to the year 2000.

**Table 5.1:** Ranking Mexican Universities 1992-2016 for STEM Sciences.

Tier 1	Tier 2	Tier 3	Tier 4
1. CINVESTAV <sup>a</sup> (2.5294)	10. UAM (2.684)	19. BUAP (3.203)	28. TEC (4.8494)
2. CICESE (2.6497)	11. UAEM (4.0295)	20. UASLP (3.5339)	29. UDG (3.9136)
3. UNAM (1.3432)	12. INECOL (4.3667)	21. CIMAV (4.7517)	30. UGTO (3.8549)
4. CINVESTAV <sup>b</sup> (2.0916)	13. IBERO (7.3052)	22. IPN (3.0192)	31. IMSS (3.7836)
5. IMP (2.4086)	14. CICY (4.641)	23. UV (4.778)	32. UNISON (5.8237)
6. ECOSUR (3.0022)	15. CIO (4.8308)	24. UADY (6.7742)	33. UMich (4.8774)
7. IPICYT (4.5405)	16. COLPOS (4.937)	25. UAEM (3.8454)	34. UAEH (4.6305)
8. CIBINOR (4.8267)	17. CIQA (6.7092)	26. UCOL (4.6986)	35. UANL (3.8814)
9. INAOE (4.3645)	18. CIAD (6.5083)	27. UAQ (6.3504)	36. TECNAM (3.2274)

<sup>a</sup> Zacatenco unit. <sup>b</sup> All units but Zacatenco. (\*) Brackets correspond to standard deviations. Full names of universities are listed in Appendix D.22.

**Table 5.2:** Ph.D. to First-Job Mexican Faculty Hiring Matrix 1992-2016 for STEM Sciences.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36
1. CINVESTAV <sup>a</sup>	0	0	2	2	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2. CICESE	0	5	2	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	0	3	1	0	0	0	0	1	0	0	0	0	0	0	0
3. UNAM	0	1	178	10	3	2	2	2	2	13	11	3	0	2	1	0	1	2	7	2	1	30	4	0	4	1	3	7	6	5	0	2	6	10	2	4
4. CINVESTAV <sup>b</sup>	0	1	9	22	1	0	0	0	0	0	1	1	0	4	0	0	0	2	4	6	2	12	1	1	1	5	0	1	4	5	2	2	3	2	5	2
5. IMP	0	0	2	0	16	0	0	0	0	0	3	0	0	1	0	0	0	0	1	1	0	5	2	0	0	0	0	1	1	0	0	0	1	0	0	0
6. ECOSUR	0	0	2	0	0	2	0	0	0	1	0	0	0	1	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0
7. IPICYT	0	0	2	0	0	0	4	0	1	0	0	0	0	0	0	0	0	0	0	0	0	3	1	0	1	1	0	0	0	1	0	0	0	0	0	0
8. CIBINOR	0	1	1	0	0	0	0	6	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	
9. INAOE	0	0	2	0	0	0	0	0	2	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	2	0	0	0	0	0	0	
10. UAM	0	0	4	1	0	0	0	0	16	0	0	0	0	0	0	1	0	0	2	2	0	5	1	0	1	0	0	0	0	0	0	2	1	0	0	0
11. UAEM	0	0	7	0	1	0	0	0	0	1	7	0	0	0	0	0	0	0	0	0	1	0	2	0	0	0	0	0	1	0	0	0	0	0	0	1
12. INECOL	0	0	1	0	0	0	0	0	0	0	4	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0	
13. IBERO	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	
14. CICY	0	0	2	3	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	1	0	0	0	2	
15. CIO	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	1	
16. COLPOS	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	
17. CIQA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	7	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	
18. CIAD	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	4	0	0	1
19. BUAP	1	0	1	1	0	1	0	0	1	0	0	0	0	0	0	0	0	0	9	1	0	2	1	0	0	0	0	0	0	0	0	0	1	1	0	0
20. UASLP	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	9	0	0	1	0	1	0	0	0	1	1	0	0	0	1	1	0
21. CIMAV	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
22. IPN	0	0	3	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	8	0	0	0	0	0	0	0	2	0	0	1	2	1
23. UV	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	1	0	0	6	0	0	0	0	0	0	0	0	0	0	0	0	0	2
24. UADY	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0
25. UAEM	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	1
26. UCOL	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	2	0	0	0	0	0	0	0	0	0	0	0
27. UAQ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	0	0	0	0	0	0	1	
28. TEC	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	0	0	0	0	0	0	0	
29. UDG	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	6	0	1	0	0	0	0	1	
30. UGTO	0	0	2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5	3	1	0	0	0	0	0	
31. IMSS	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5	0	0	1	0	
32. UNISON	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	6	0	0	0	0	
33. UMich	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	0	0	0	
34. UAEH	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	
35. UANL	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	
36. TECNAM	0	0	1	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0

Entries in the matrix correspond to the number of Ph.D. graduates from university  $i$  hired by university  $j$  for their first-job. Full names of universities are listed in Appendix D.22.

**Table 5.3:** Prestige movements per university prestige 1992-2016. The table presents how many scholars in each university tier moved the prestige hierarchy considering Ph.D. institutions (top) and first job (bottom).

	down	stay	up	tot %
Ph.D. Tier 1	304	283	17	68 %
Ph.D. Tier 2	54	49	22	14 %
Ph.D. Tier 3	38	48	12	10 %
Ph.D. Tier 4	10	47	14	8 %
Job Tier 1	49	283	46	42 %
Job Tier 2	64	49	13	14 %
Job Tier 3	144	48	2	22 %
Job Tier 4	149	47	4	22 %

## 5.6 Matched Pairs Analysis

In this section, I examine how movements in the prestige hierarchy in the transition from the Ph.D. the first-job affects scholars academic performance. I test it by using a quasi-experimental technique based on matching, which was developed in the previous chapter and this section summarises it briefly.

The technique compares the academic performance of matched pairs of scholars in the three groups, *Up*, *Down* and *Stay* matching individual characteristics and placement (or hiring) institution. When individuals are paired holding constant their Ph.D. institution, I test how changes in prestige relate to academic performance irrespective of training. Similarly, the comparison done matching scholars by their first-job examines how prestige movements affect scholars performance for agents with the same first academic job. Since early-career changes in prestige might impact academic performance differently as time passes, I compare the performance of matched pairs of scholar between groups in the *short* (up to 2 years), *medium* (3-5 years) and *long-run* (6-25 years) after graduation.

The method combines exact and random matching. It starts from the smaller group and for each individual finds at least one match in the other group. If there is only one possible match the pair is formed<sup>5</sup>. Otherwise, if in the larger group there is a set of individuals with the same characteristics, that can match the one in the smaller group, I form the pair drawing one at random in that set. I apply a re-sampling technique to the matching methodology described above to reduce small sample biases as it follows. For each comparison *Up* vs. *Stay*, *Down* vs. *Stay*, and *Up* vs. *Down* I generate  $n = 10000$  bootstrap samples of the group on the left-hand side (the smaller) of its same size  $s$ . For each of the 10000 samples of size  $s$ , I create matched pairs of scholars matching on *gender*, *age*, *discipline*, *graduation year* and Ph.D. (or first job) university. In order to compare their performance, in each sample I estimate the proportion of pairs,  $p^* = (p_1, p_2, \dots, p_n)$ , in which one group  $g^\alpha$  have higher performance than the other (group)  $g^\beta$ . Such as

$$p_n^\alpha = s^{-1} \sum_{i=1}^s I(g_i^\beta < g_i^\alpha) \quad (5.4)$$

$$I(g_i^\beta < g_i^\alpha) = \begin{cases} 1, & \text{if } g_i^\beta < g_i^\alpha \\ 0, & \text{Otherwise.} \end{cases}$$

Where academic performance  $g$  is the individual NSR research rating.<sup>6</sup> For each

<sup>5</sup>Given how prestige operates it implies that the sizes of the groups *Up*, *Down*, *Stay* are different and are likely to have sizes ordered in this way  $Stay > Down > Up$ . Thus, in the comparison *Down* Vs. *Stay* *Down* is smaller than *Stay* and it can be considered the “treatment” group. That is the pairs are done starting the search from the individuals in *Down* looking for a match in *Stay*. In contrast, in the comparison *Up* vs. *Down* *Up* is smaller than *Down*; So the pairs will be formed starting from elements in *Up* and looking for matches in *Down*

<sup>6</sup>NSR rating is one of 5 ordered categories

group in the comparisons *Up vs. Stay*, *Down vs. Stay* and *Up vs. Down*, I estimate the two  $p^*$  and construct their  $F(p^*)$  cumulative empirical distribution function (CEDF). To assess the performance of one group over the other I test for first order stochastic dominance (Levy, 1992). This test implies higher performance of  $g^\alpha$  over  $g^\beta$  if  $F(p^\alpha) \leq F(p^\beta)$  for all  $p^*$ .<sup>7</sup> I compare the two CEDFs running a *two-sided* and a *one-sided* Kolmogorov-Smirnov test (*KS test*).

The null hypothesis of a *two-sided* test is  $H_{01} : F(p^\alpha) = F(p^\beta)$  — the two CEDF are drawn from the same distribution. Rejecting the null hypothesis  $H_{01}$  implies that the academic performance is statistically different between the two groups. The null hypothesis of the *one-sided* test is  $H_{02} : F(p^\alpha) \geq F(p^\beta)$ . Rejecting the null hypothesis implies that  $F(p^\alpha)$  stochastically dominates  $F(p^\beta)$ , in other words, that the increase in academic performance associated with a change of prestige from group  $g^\alpha$  is statistically different and greater than  $g^\beta$ .

## 5.7 Matched Pairs Results

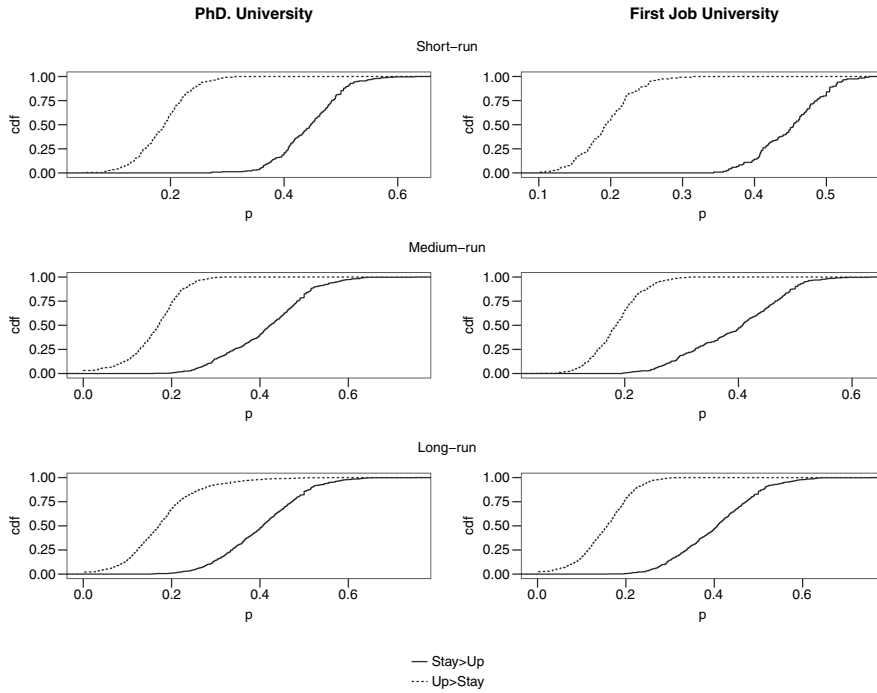
In this section, I compare scholars performance of *Up vs. Stay*, *Down vs. Stay*, and *Up vs. Down* in the short, medium and long-run. In particular, I examine the CEDF of the proportion in which one group received a higher NSR rating than the other. *Up*, *Down*, and *Stay* represent prestige changes from the Ph.D. to the first-job where prestige is measured with the interactive ranking with a moving time window of 3 years.<sup>8 9</sup> Results of the KS-tests of  $H_{01}$  and  $H_{02}$  in Appendix D.26 show for every comparison that the CEDFs are different and one group stochastically dominates the other. I find consistent results both using the dynamic and the static prestige ranking, where the latter is in Appendix D.25. The similarities between the two results do not necessarily suggest that the distribution of prestige is stationary but rather that the association between changes in prestige during early career and academic performance are consistent, irrespective of changes in the hierarchy.

Figure 5.1 compares the NSR research performance of matched pairs of scholars who *Stay* and move *Up* the hierarchy. In all figures, scholars match if they have the same gender, age, discipline, and graduation year. Additionally, figures on the left match scholars with the same Ph.D. while those on the right-hand side match those with the first-job institution. The matching procedure allows us to compare scholar with same Ph.D. (or first-job) and characteristics but experiencing different prestige movements. In both cases, the CEDF of *Stay* > *Up* (solid lines), is located below that of *Up* > *Stay* (dashed lines) implying that the *Stay* group stochastically dominates the *Up* group. The implication of the results is the following. On the one hand, looking at scholars with the same (Ph.D.) training (left-plots) I find that those hired internally have on average a better research performance than those experiencing up-

<sup>7</sup>Graphically this is inspected if  $F(p^\alpha)$  lies below and to the right of  $F(p^\beta)$ .

<sup>8</sup>Details on the dynamic rank computation are in section 5.4.

<sup>9</sup>I execute the algorithm with a time window of  $\Delta = 3$  and run robustness checks using other time ranges ( $\Delta = 5$  and  $\Delta = 8$ ) and the results are consistent.

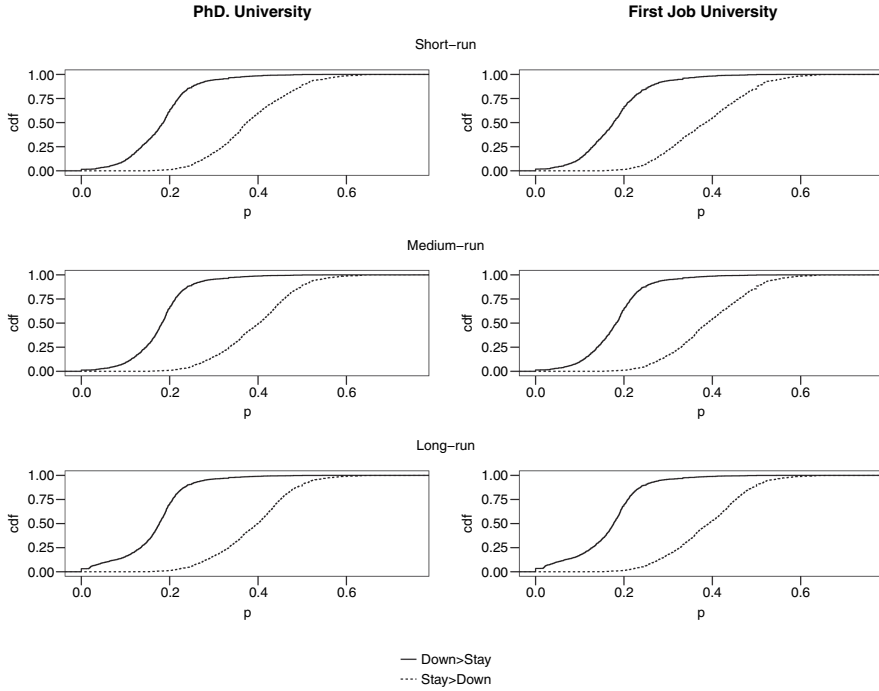


**Figure 5.1:** Up versus Stay comparison. The solid curves are CEDFs of the proportion of pairs in which  $R_{Stay} > R_{Up}$ . Dotted curves are CEDFs for  $R_{Up} > R_{Stay}$ . Pairs matched by gender, age, discipline, graduation years, and same Ph.D. university (left), or same first-job university (right). From top to bottom: short-run (Up to 2 years), medium-run (3-5 years), and long-run (6-25 years) after Ph.D. graduation.

ward prestige mobility (hired into a university more prestigious than their Ph.D.). On the other hand, comparing scholars with the same first-job (right-plots) but different training (Ph.D.) I find that internal hired perform better than those coming from less prestigious Ph.Ds. (Up). These results suggest that scholars who manage to secure positions at their faculty after graduation demonstrate higher NSR levels of performance than those who migrate to upper ranked institutions.

Results for the comparison between the *Down* and *Stay* groups are in figure 5.2. In this case, I compare scholars who take academic positions in their faculties after graduation and Ph.D. graduates experiencing downward prestige mobility. Results are the same matching the pairs on the Ph.D. or the first-job institution — the CEDF  $Stay > Down$  stochastically dominates  $Down > Stay$ . In line with previous results, plots on the left-hand side show that scholars with the same Ph.D. training moving (down) to a less prestigious institution in their first-job tend to have a lower NSR rating than those hired internally (stay). Similarly, plots on the right-hand side compare scholars with the same first-job and indicate that those hired internally (stay) have a higher performance than those moving down the hierarchy.

The last comparison in Figure 5.3 examines performance differences between



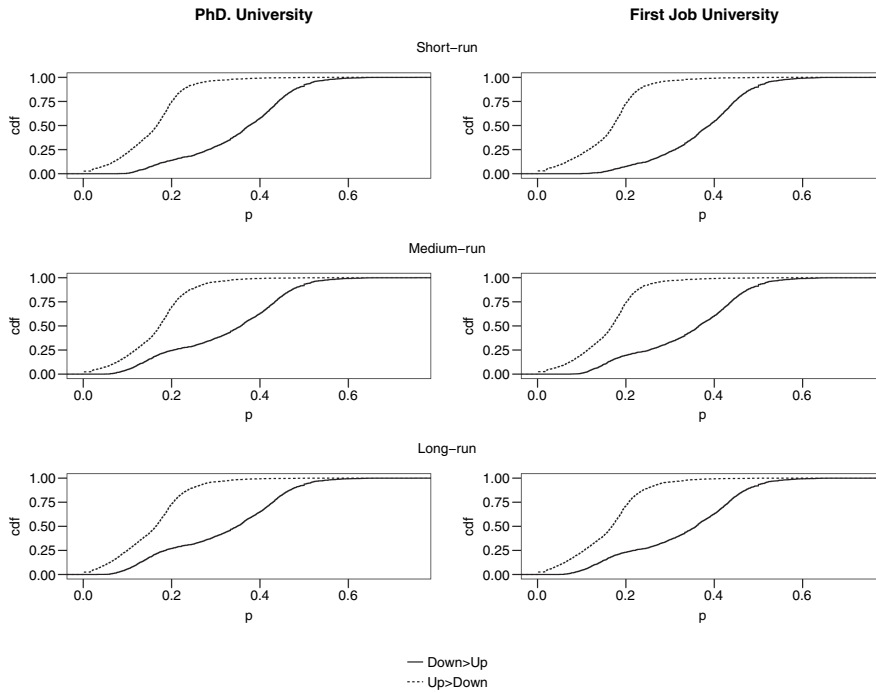
**Figure 5.2:** Down versus Stay comparison. The solid curves are CEDFs of the proportion of pairs in which  $R_{Down} > R_{Stay}$ . Dotted curves are CEDFs for  $R_{Stay} > R_{Down}$ . Pairs matched by gender, age, discipline, graduation years, and same Ph.D. university (left), or same first-job university (right). From top to bottom: short-run (Up to 2 years), medium-run (3-5 years), and long-run (6-25 years) after Ph.D. graduation.

scholars who experience upward and downward prestige mobility. Results show that the CEDF of the proportion of pairs of scholars where the performance of  $Down > Up$  stochastically dominates  $Up > Down$ . The stochastic dominance of one over the other implies that scholars who experience downward prestige mobility sustain higher performance over their career than those experiencing upward mobility in their early career. These results are consistent both matching pairs keeping fixed the (Ph.D.) training (left-plots) or the first-job (right-plots) institution. In the first case, comparing scholars with the same (Ph.D.) training, I find that those moving down the hierarchy have higher performance than those moving up to more prestigious first job institutions. The second, pairing scholars with the same first job but different Ph.Ds. institution, I find that those coming from more prestigious Ph.Ds. (down) have a higher NSI rating on average than those moving up from less prestigious Ph.D. institutions.

These results seem counter-intuitive at first glance since most studies have associated upward ranking mobility with higher academic performance.<sup>10</sup> In particular,

<sup>10</sup>Moreover, I should consider that those that start high in the hierarchy have few options to move up the ranking and they are more likely to move down, while the opposite is true for those low in the prestige hier-





**Figure 5.3:** Down versus Stay comparison. The solid curves are CDFs of the proportion of pairs in which  $R_{Down} > R_{Stay}$ . Dotted curves are CDFs for  $R_{Stay} > R_{Down}$ . Pairs matched by gender, age, discipline, graduation years, and same Ph.D. university (left), or same first-job university (right). From top to bottom: short-run (Up to 2 years), medium-run (3-5 years), and long-run (6-25 years) after Ph.D. graduation.

the first result of a negative impact of upward prestige mobility comparing scholar with the same training contradicts the previous results of Chan et al. (2002). However, their analysis is slightly different. They use a longitudinal analysis in one sub-field of economics and their sample is limited to scholars with publications in 16 top journals in finance. In contrast to us, they find that scholars who experience upward prestige mobility publish twice as many as their colleagues.

What is most interesting is that what I found is a pattern through the career of scholars and for both dynamic (Figure 5.3) and static ranking estimations (Appendix D.25). Nevertheless, these findings require further research that I discuss in section 5.8.

### 5.7.1 Robustness check

In this sub-section I provide as a robustness check a regression analysis aimed at offering complementary results. In particular, I examine whether an alternative ap-

---

archy. This implies that those moving up (down) are more likely to have Ph.Ds. in less (more) prestigious institutions.

**Table 5.4:** Ordered logistic regressions: The impact of prestige changes (Up, Down, Stay) on academic performance. The dependent variable is the individual NSR research rating in 5 ordered categories (see Appendix D 24 for details). The main dependent variables are the dummies Up, Down and Stay. Where Down is the baseline category.

	Model 1	Model 2	Model 3
PhD-Top-10	0.40860 (0.23103)		0.52041* (0.23784)
PhD-Top-20	-0.04998 (0.19665)		-0.07218 (0.19566)
PhD-Top-30	0.09305 (0.15779)		0.07570 (0.15924)
<i>Up</i>		0.01439 (0.21731)	0.19335 (0.23386)
<i>Stay</i>		0.30306* (0.11890)	0.34912** (0.12012)
Male	0.37390** (0.11388)	0.40277*** (0.11260)	0.40828*** (0.11415)
Age	0.01767 (0.01365)	0.01921 (0.01383)	0.01764 (0.01385)
Discipline II	-0.19528 (0.15857)	-0.17127 (0.15540)	-0.17966 (0.15630)
Discipline III	-0.45132 (0.31462)	-0.53938 (0.30761)	-0.46020 (0.31410)
Discipline VI	-0.56758** (0.21150)	-0.57312** (0.20922)	-0.55852** (0.21395)
Discipline VII	0.05683 (0.17391)	0.04987 (0.17424)	0.08791 (0.17439)
AIC	9982.40999	9972.24112	9954.94310
BIC	10247.63255	10230.99483	10233.10335
Log Likelihood	-4950.20500	-4946.12056	-4934.47155
Deviance	9900.40999	9892.24112	9868.94310
Num. obs.	4764	4764	4764

Clustered standard errors (individual): \*\*\* $p < 0.001$ , \*\* $p < 0.01$ , \* $p < 0.05$

proach confirms the previous results. Thus, I run *Ordinal Logistic Regression* models using as the dependent variable the research performance proxies by the 5 ordered NSR rating categories (See Section 5.3) where the main explanatory variables are prestige changes dummies Up and Stay where Down is the baseline category. I use as additional control variables the prestige of the Ph.D. institution, age, gender, year of Ph.D. graduation and discipline.

Results of models 2 and 3 in table 5.4 show a positive and significant association between the likelihood of a high NSR rating and scholars experiencing internal hiring (Stay) in comparison to those who move down. The lack of a statistically significant relationship between upward mobility and performance (when Up is compared to Down) also suggest that scholars hired in their Ph.D granting institution perform

better than those experiencing upward prestige mobility.

## 5.8 Discussion

The main findings underline a complex relation between prestige movements and performance. In particular I find a positive association between internal hiring and research performance in comparison to scholars moving in the hierarchy. Moreover, I find that scholars experiencing downward prestige mobility show higher performance compared with colleagues with similar characteristics moving upward in the prestige hierarchy. A similar phenomenon can be observationally equivalent to various mechanisms that require further research. Below I provide possible explanations and findings in the past literature. In particular, I focus of the BFLPE, peer effects, and the “*what does not kill me makes me stronger*” effect.

Psychology research in self-cognition underlines the importance of academic self-concept in educational settings. The notion of academic self-concept relates to the beliefs an individual has on her/his academic aptitude and academic achievements. This comes to existence because the view that agents have on themselves is not always accurate. For example, academic self-concept of an individual is accurate when is aligned with her/his (objective) outcomes, such as scholastic aptitude, intelligence test scores and actual academic performance. However, in many cases individuals view of themselves may be odd. This often manifests when an individual compares herself/himself with peers. In similar cases, the social comparison is an assessment that may impact for better or worse on individual performance. This hypothesis is called the BFLPE, where the academic self-concept of a student is determined both on her/his academic achievement levels and on the average achievement levels of the other students in her/his school.<sup>11</sup> According to this hypothesis a student will have a lower academic self-concept in an academically selective school, where the average achievement of peers is high, than in a non-selective one (Werts and Watley, 1969; Astin, 1969; Marsh and Hau, 2003; Marsh et al., 2008; Salchegger, 2016; Rosman et al., 2020; Von Keyserlingk et al., 2020). In terms of performance, the BFLPE hypothesis predicts that a lower level of academic self-concept is likely to be associated to a lower performance and to potentially undermine students future achievements (Marsh and Hau, 2003).

Many empirical results confirm the BFLPE and found it to be strong in different cultural settings. For example, in a large cross-country comparison Marsh and Hau (2003) find a strong support for the BFLPE. The study samples 3849 schools in 26 countries, looking a total of 103558 students of age 15 for the year 2000. They use as dependent variable the individual self-concept measured in the Cross Curriculum

---

<sup>11</sup> Still is unclear when academic self-concept first manifest in children. For Tiedemann (2000) academic self-concept develops in children from age 3 to 5 because of family and early educators’ influences. In contrast, more recent studies suggest that academic self-concept develops much later between the age of 7 and 11 years old (Leflot et al., 2010; Rubie-Davies, 2006).

Competencies survey. While their main independent variable is the average PISA test score of the student's school, which it proxies how much the school is selective. All countries, with the exception of Korea and Hungary, display a negative and significant coefficient of school average score supporting the BFLPE. However, very little research focuses on higher education settings and most of the empirical results related to the BFLPE has been conducted in children or adolescents at school age (Salchegger, 2016). Indeed, beside the evidence in students of young age, a similar mechanism that explains how social comparison influences the beliefs and, consequentially, the individual achievement can be relevant also among adults. Recent advances study the BFLPE in young adults. Rosman et al. (2020) test the BFLPE in a sample of 115 first-year undergraduate psychology students at the Leibniz Institute in Germany. They find no support for the BFLPE, in their sample entering the university setting does not modify students' academic self-concept. However, their sample is small and not representative of students in higher education. In contrast, in a larger and more representative study, focusing on 2182 German students in the transition from high school to universities, Von Keyserlingk et al. (2020) find strong evidence for the BFLPE.

Adults potentially are more used to competition and the BFLPE may be less relevant for them. However, becoming a scientist in academia has career phases where the social comparison might matter. The transition from the Ph.D. to the first academic job is the ideal moment to look at similar mechanisms, especially when individuals change environment.<sup>12</sup> Indeed, when individuals change environment, or advance in their career, the group of peers also changes. For example, after Ph.D. graduation scholars can be hired internally or move to different institutions. Thus, scholar mobility implies that individuals who change affiliation consequentially change their peers and their place in the prestige hierarchy of the university system. Similar changes in social comparison mechanisms could affect individual performance. BFLPE is one possible mechanism behind my findings in this chapter. In my opinion, the BFLPE should be further investigated in university settings in the future.

In an academic setting other mechanisms may produce outcomes similar to the BFLPE. In more general terms, learning and social comparison can be coupled and generate various peer effects in an academic environment that in turn influence performance (Hottenrott and Lawson, 2017). Most of the literature, looking at scholars in academia, focuses on how social comparison generates positive peer effects. However, past research underlines that, depending on the level of the competition, peer effect can also be negative (Stapel and Koomen, 2005). This also refers to the unit of analysis, for example at the group level the peer effect can be positive and at the same time it can be negative looking at the individual level. This will happen when the group benefits from the contribution of a new member but, at the same time, the group has a negative effect on the newcomer's performance.

---

<sup>12</sup>In the literature the BFLPE is predicted to be stronger when individual change environment. See the case study in (Marsh, 1991).

For example, Slavova et al. (2016) study the peer effect that social comparison has at the group level. They examine whether hiring a new scientist affects the scientific performance of the incumbents in the hiring department. Their empirical focus is a sample of 94 U.S. chemical engineering departments between 1996 and 2004. They find that a new hire generates a positive peer effect in the performance of the department. However, their results show that only colleagues with a recent tenure were affected. As a limitation their analysis does not consider what is the impact of changing department peers for the newcomer. Related to this, they suggest that peer effects can be positive or negative depending on scientific specialization and the level of competition/collaboration within the department. A department more specialized and with more internal collaborations will be more likely to integrate the new scientist and experience positive peer effects compared to a less specialized and more competitive one.<sup>13</sup>

The way in which people are integrated or promoted in a new workplace can affect individual self-esteem and academic self-concept as well. In general, the literature associates high self-esteem and self-concept to high career outcomes. However, there are cases where this association appears to be negative. Other research shows that individuals with high self-esteem have negative outcomes in terms of work effort and performance (Whelpley and McDaniel, 2016; Sherf and Morrison, 2019; Li et al., 2020). Weiss and Knight (1980) find that individuals with high self-esteem exert less effort in searching for information and have lower performance in comparison with people with lower levels of self-esteem.

Moreover, in academia a rejection can have a negative impact on self-esteem and academic self-concept, however, it can also generate positive effects on performance. The latter relates with the mechanism “*what does not kill me makes me stronger*”. In a recent contribution Wang et al. (2019) compare publication and citation records of 561 narrow wins and 623 near misses scientists who applied to the NIH grant<sup>14</sup>. Overall they find that those rejected have 10% chance of disappearing permanently from the NIH granting system. Yet, despite an early setback, individuals with near misses proposal (who persevere) systematically outperform those with narrow wins in the longer run. This result could be consistent to a BFLPE in early career scientists, where an initial promotion or confirmation of abilities may be counter-productive for future performance.

---

<sup>13</sup>This suggests that social comparison creates peer effects that are context specific. Indeed, the internal level of competition affects the behaviour of both the group of incumbents and the newcomer. When resources are scarce, the level of competition to access them increases and in extreme cases a newcomer can be perceived by the group as a threat. In this case the effect of the social comparison may negatively affect the performance. An extreme case of thereof are bullying episodes or misbehaviour in academia (McKay et al., 2008; Keashly and Neuman, 2010; Giorgi, 2012). See Henning et al. (2017) for a review on the topic. However, most of this research is limited to voluntary surveys with little representativeness and mostly on administrative staff at universities.

<sup>14</sup>NIH (National Institutes of Health) is the largest public funder of biomedical research in the world

In light of the mechanisms mentioned above, we can interpret the results of the Mexican Ph.D. job market as follows. Scholars who move down might think they are “big shots” relative to their new peers and this is beneficial to their performance; or their new colleagues might think they have made a catch and give them more resources; or an initial “failure” - moving down in the hierarchy - might lead to efforts to “regain previous prestige”.

All of which would be consistent with the apparently paradoxical result in this chapter and of essential consideration for future research and policy.

## 5.9 Conclusion

This chapter is the first studying how prestige stratification affects scientific performance of early career scientists in the Mexican higher education system. The majority of comparable analyses looked at university systems in developed economies especially in North America, where mobility after the Ph.D. tends to be high and systems more integrated. In the U.S. for example there are often hiring practices that prevent universities from hiring their own graduates immediately after the Ph.D. Studying these mechanisms in less mature settings with higher resource constraints has policy implication because the gap between prestige and performance can be larger than other settings.

The findings in general suggest that there is a negative relation between mobility during early-career and academic performance. Moreover, when I decompose mobility looking at prestige differentials between Ph.D. and first-job institution, I find that scholars who *Stay* or move *Down* the hierarchy remain mostly in first-tier (top 10) institutions.<sup>15</sup> Results of the matched pairs analysis provide evidence of the same association of prestige movements and performance in the short-, medium- and long-run. Further comparing those moving up with those moving down the hierarchy I find that those moving down have sustained higher performance than those moving up.

The reasons why promising scholars experiencing upward prestige mobility have lower performance than their colleagues (with the same Ph.D.) that stay or move downwards requires further investigation as highlighted in the previous section. A similar finding is somewhat reminiscent of the work by Robert Frank in the 1980s regarding “choosing the right pond” in which the author found that some highly productive individuals choose to work in low prestige environments where they gain status and perform well by being at the top (Frank, 1985).

The counter intuitive result of this chapter requires further scrutiny. In particular, it should be interpreted considering that Mexico is potentially subjected to both a *brain-drain* and a *brain-gain* phenomenon. Indeed, this chapter considers Mexico as a close system but many Mexican researchers migrate to and from the US to continue

---

<sup>15</sup>Those moving *Up* the hierarchy, get their Ph.D. degree mostly from second-tier (bottom 30) institutions but move to first-tier institutions with their first job. This is an expected result in the light that those graduating from prestigious universities have fewer possibilities to move higher in the hierarchy.

their research. In the case of Mexico, there are a considerable number of Mexican individuals who receive their Ph.D. training in the US and then they return to Mexico. For example, Finn (2011) estimates that only 40% of Mexican Ph.D. graduates, trained in the US, are hired in a US university.<sup>16</sup> This suggests that at least some of them will return to Mexico and they are likely to receive a position in a Mexican university. This might increase the competition between young scholars and those trained in the US might receive more resources or continuing their collaborations with US scholars.

Similarly to the higher education system in the U.S. and other developed economies, I find a large stratification in the Mexican university system but low mobility (around 50% of Ph.D. graduates are hired by their faculty), with a few prestigious institutions (around 10) producing the majority of Ph.D. graduates that are subsequently mostly hired in these same institutions. A high concentration of prominent scholars in a few academic institutions reveals large inequalities in the distribution of prestige. On the one hand, the stratification of higher education could promote higher levels of specialization with a targeted allocation of resources. On the other hand, it can also reveal a structural problem in the science system, a “lock-in”, where researchers are trained and hired by elite institutions and flows of knowledge are reduced throughout the national science system. In the case of Mexico a structural “lock-in” could be additionally reinforced by the negative association between mobility (upward prestige mobility in particular) and performance.

---

<sup>16</sup>See <https://orise.ora.gov/stem/reports/stay-rates-foreign-doctorate-recipients-2011.pdf>; Last access December 2020.

## Chapter 6

# Conclusion

This thesis proposes new methods to study social transformations by examining the entry process in a workplace and it uses the empirical focus of the South African and Mexican Academia.

As mentioned in the previous chapters, the entry process in the labour market and social transformations are tightly linked. The entry process changes the demographic composition of a workplace and is a main channel to social transformations. In general, the entry process follows systemic and individual level mechanisms.

Systemic mechanisms influencing the entry process are the interaction of the demographic groups in a population, the relation between the old and the new generations, the level of homophily, Matthew effects and the connections between institutions of different status.

The mechanisms at the individual level that play a role in the entry process are the individual output, the social connections of the individuals, the training, the relation between the trainee and his/her advisor, and mobility.

I examined individual and system level mechanisms. Indeed, this thesis relied on a theoretical framework which connects the economic literature (focusing on individual mechanisms) and the sociological literature (which concentrates on systemic characteristics) to explain the entry process in the workplace. In the preceding chapters, I developed new empirical methods to integrate the two perspectives.

The chapters of this work focused on career stages which are crucial in the entry process. Indeed, the entry process has many dynamics and potential bottlenecks which can enhance or inhibit social transformations. In the case of academia, the main steps of the entry process are: education (especially at the master level), doctoral work and training, and the first job after the doctorate. The general framework of the thesis analysed these three periods of a career in academia.

In Chapter 2, I studied the relations of training during education, which in academia corresponds to student thesis supervision. Related to this, I focused at the whole sam-



ple of students (and their advisors) before any selection takes place. I looked beyond labour market outcomes, examining homophily in the formation of the relation between the student and their advisors.

In Chapter 3, I looked at training more closely. I focused on a smaller sample representing the promising students likely to make an entire career in academia, the Ph.D. students. I studied individual productivity during the doctorate including the supervisor characteristics.

In Chapter 4 and 5, I looked at the first job after graduation. I focused on the transition from the Ph.D. to a first job; including those who obtain a first job in academia within five years after graduation. I examined the link between Ph.D. to first job hiring flows, institutional prestige and research quality.

Beyond the general framework, each chapter examined systemic and individual mechanisms of the entry process in a workplace. The chapters of this thesis provided interesting results and new empirical methods.

In Chapter 2, I found that the formation of student-supervisor relation can operate as a first bottleneck preventing social transformations.

I developed a new methodology to separate choice homophily from that induced by the system which solves causal issues creating a null model of type-blind tie formation that incorporates network structures and individual characteristics (i.e. supervision capacity, popularity, department norms). Using those results, I examined how population groups relate one other. I developed a model with interacting populations and estimated homophilous preferences in the demographic groups using aggregate department level data of student, supervisor, and tie composition. The methodology can be used for monitoring social transformations within organisations and it is suitable in the presence of sparse register data. I applied the described methods to register data of student-advisor relations at South African universities between 1973 and 2014.

My results underline that in the university system both race-based and gender-based homophily influence tie formation. The observed racial (gender) homophily in the system arises from institutional constraints, network structures, and choices made by individuals. In particular, half of the observed homophily is induced by the system through institutional constraints and network structures and half is driven by choices. The latter originates primarily in new members of the former dominant group (white/male students). However, when the sample of students is restricted to promising students who enter academia, white professors also display choice homophily. Further, I found that during a transition phase of social transformations homophily can mechanically increase even if there are no behavioural changes among individuals, under certain circumstances. In terms of policy recommendations, Chapter 2 underlines that to achieve sustainable social transformations a demographic change is not enough and that structural constraints and behavioural aspects should be tackled together. In particular, white (male) students and their advisors is a critical target group for intervention.

In Chapter 3, I focussed on the individual mechanisms related to training, testing the link between homophilous/heterophilous (same-/cross- type) ties and productivity. In particular, I studied whether student-advisor gender and racial couple affects student's doctoral productivity. I used an econometric technique, employing a panel data analysis with a quantile regression approach. This approach studies the link between homophilous/heterophilous ties and productivity looking at different student productivity-profiles. I found that on average female students publish less than males. This difference exists for female students working with a male advisor (heterophilous ties) and not those working with a female supervisor (homophilous ties). But looking at this link for different student productivity-profiles I found the following. In the whole sample, female students with a high (or low) productivity-profile studying with female advisors are as productive as male students with a high (or low) productivity-profile studying with male advisors. Decomposing the joint effect of gender and race I found different results looking at same- and cross-racial supervision. For same-race supervision, the gap in average performance between female students and male students is mainly driven by a gap in the right hand tails of the productivity distribution (students with high productivity profiles). In contrast, cross-racial supervisions display a smaller gender gap that is u-shaped over student productivity.

Chapter 3 underlines the role that female advisors have for female students. Further, it shows the crucial role that cross-racial ties have in reducing gender productivity differences. Moreover, the chapter points out that same-(cross-) racial ties have a heterogeneous effect on student productivity depending on her/his productivity-profile. In terms of policy advice, the evidence from the chapter shows that promoting cross-racial ties and female supervision have potential in decreasing the gender gap.

Chapters 4 and 5 focused on the micro and macro mechanisms of the Ph.D. job market, I studied the link between hiring, university prestige, and research quality. I developed a technique to rank university prestige and used it to test whether prestige movements affect scholar research quality. To achieve this, I developed a new quasi-experimental method, based on matching. I compared the quality performance of matched pairs of individuals with the same training (or same job) but different movements in the prestige hierarchy.

I found that the university system has a hierarchical prestige ordering with specific hiring flows. In this stratified hierarchy, blacks and females are disadvantaged — they tend to experience more mobility and to move down in the prestige hierarchy when hired in their first job in academia. Overall, inertia has a positive role in the system: those who make large movements in terms of prestige have lower research ratings than those who are hired by the same university granting the Ph.D.. This inertia also affects the long run (20 years after Ph.D. graduation). The results in the Chapters suggest that the hierarchical stratification of the university system is a constraint to social transformation because it systematically disadvantages upward (prestige) mobility (of black and female Ph.D.s in particular) with effects also in the long-run. In terms of

policy, any intervention aimed at counteracting this hierarchical stratification allowing a fruitful exchange between institutions will promote social transformations. Further, developing incentives to allow black and female students to stay after Ph.D. and built their research capacity before moving institution could also be beneficial.

Besides the policy advices that this research have raised, the work have some limitations that will require a further scrutiny. A limitation of the work is that it looks at social transformations in the entry process in academia. But the different levels of social transformations within an organisation could act differently on individuals depending on the period in the career. Further, the hierarchical structure of the university system is endogenous and created by multiple channels of interaction between agents, departments, and institutions.

Future expansion of this work will study social transformations looking at agents longitudinally over the career and incorporating the endogenous nature of the hierarchical structure of the university system looking at multiple channels of interaction.

As a general remark, this research concentrates in academia. Academia is a knowledge-intensive and desirable profession, but above all is the "temple" of knowledge and for centuries exclusively for white males. Women and people of colour for centuries were prevented from actively participating in the higher education system. Indeed, both the South African Apartheid and segregation in the U.S. were designed to impede people of colour to access quality education. Access to quality education has always been at the top of the agenda of civil rights moments. In 1906, The U.S. scholar William E. B. Du Bois, the first African American to hold a doctorate, included education as the fifth priority for the fight for civil rights. In the second annual meeting at Harpers Ferry in West Virginia, he used these words:

"[...] Fifth. We want our children educated. The school system in the country districts of the South is a disgrace, and in few towns and cities are the Black schools what they ought to be. We want the national government to step in and wipe out illiteracy in the South. Either the United States will destroy ignorance, or ignorance will destroy the United States. And when we call for education we mean real education. We believe in work. We ourselves are workers, but work is not necessarily education. Education is the development of power and ideal. We want our children trained as intelligent human beings should be, and we will fight for all time against any proposal to educate black boys and girls simply as servants and underlings, or simply for the use of other people. They have a right to know, to think, to aspire. These are some of the chief things which we want. How shall we get them? By voting where we may vote, by persistent, unceasing agitation, by hammering at the truth, by sacrifice and work." BlackPast, B. (2010, July 26) (1906) W.E.B. Du Bois, "Men of Niagara". Retrieved from <https://www.blackpast.org/africa>

n-american-history/1906-w-e-b-dubois-men-niagara/

For the primary cultural and intellectual role of higher education and its role in a society social transformations at universities are essential. However, this work underlines how the entry process is tightly related to social transformations and how a lack of transformations in prestigious positions (universities and academic ranks) might slow down transformations at lower levels. In general terms, this work is aimed at examining how systemic mechanisms work for different demographic groups.

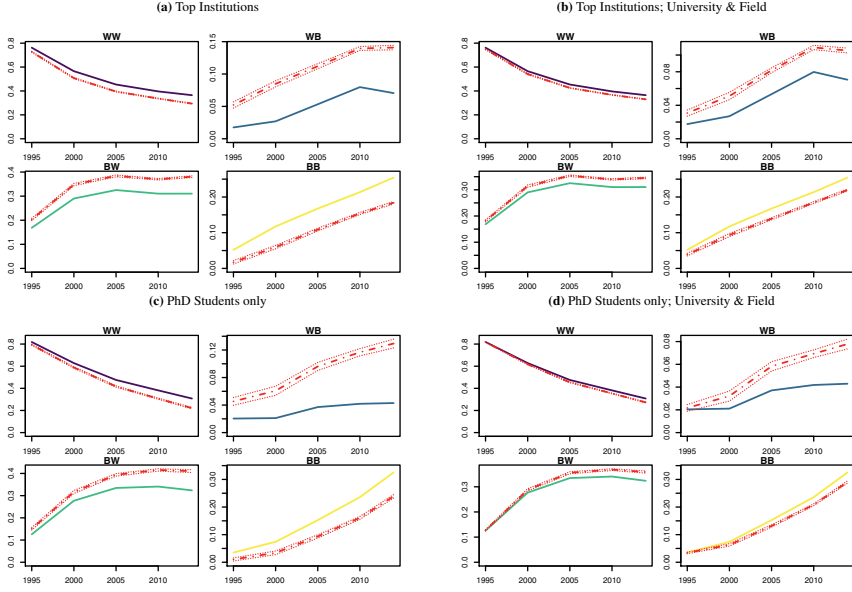
In the world, excluded groups are everywhere, rich and poor economies struggle for inclusion and in many cases, fail to provide equal opportunities for all. In order to achieve substantial and sustainable social transformations there is the need to empower those groups. Here, I focus at one of the main channel: their inclusion in the labour market. However, as a general lesson, my work shows how systemic mechanisms coupled with individual incentives produce inertia in the workplace. In particular, inertia arises from choice for association within and between individuals and institutions. The inertia acts at maintaining group and power relations unchanged, to the disadvantage of under-represented groups. Besides this disheartening result, this work also highlights a positive lesson. This inertia it is not permanent and systems contain the seed for the social change. For example, the inertia of homophilious relations coupled with the subsequent generations of workers might produce an increasing aggregate level of homophily in the system creating tensions over resource allocation. This tension eventually can destabilise the status-quo and accelerate changes. Moreover, at the individual level my findings highlight that when agents overcome the inertia of systemic mechanisms creating (or searching for) uncommon relations they perform at the highest levels underlining the aggregate gain of inclusion.

This work suggests an inherent truth about development and more generally about how societies can or cannot change. Transformations follow a non linear path traversing periods of lights and shadows. Moreover, transformations are not unidirectional and not uniformly distributed. Indeed transformations are often localised and concentrated over time and space. The concentration of transformations in particular loci is intrinsically related to the succession of different generations and the inertia produced by systemic mechanisms. The two produce tension over the accumulation of resources and are (pre-)condition for a societal change to germinate. Any process of change often implies a collective effort and painful transitions that need constant monitoring, targeting, and activism.

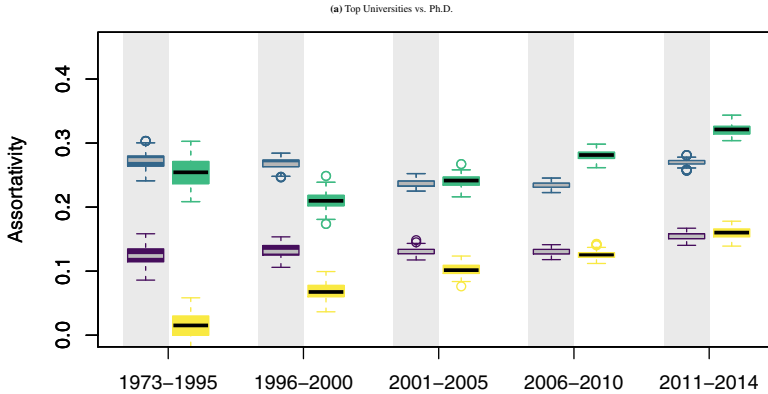


## A.1 Permutation Test results, Choice and Induced Homophily

**Figure 1:** Permutation test for student supervisor data 1973-2014 for top institutions and Phd students. The permutation is done keeping fixed the number of ties and the supervisor and permuting the students. Permutation without constraint (left) and with institutional constraints (right). Permutation is repeated 100 times for each of time period. The Observed proportions of racial student-supervisor relations (solid lines) are plotted versus the results of the permutation test (dashed lines) with respective two standard deviation from the mean. As top institutions we include: Cape Town; Pretoria KwaZuluNatal; Stellenbosch; Rhodes; Witwatersrand; and Western Cape University;"



**Figure 2:** Induced and Choice Homophily comparison looking racial ties in student supervisor relations 1973-2014. We compare results of assortativity of two sub-samples: Top Universities vs. Ph.D. Sub-sample. The panel contains 4 series: aggregated (upper series) versus department level (lower series) crossed with Top Universities (white background) versus Ph.D. (grey background). Upper series are system-level permutations, including both induced and choice homophily; lower series are department-level permutations, excluding much of the induced homophily. Each box plot represents 100 permutations and associated assortativity calculation, as described in section 2.4.2, and equation 2.4.



## A.2 List of Institutions and Fields in the data

For South African Institutions we use post 2004 merger names listed below: University Stellenbosch, University Cape Town, University Pretoria, University Fort Hare, University Western Cape, University Free State, University KwaZulu Natal, University Johannesburg, University Limpopo, Durban Institute Technology, Tshwane University Technology, Nelson Mandela Metropolitan University, Rhodes University, University Venda, Cape Peninsula University Technology, University South Africa, University Zululand, North West University, Walter Sisulu University, Vaal University Technology, University Witwatersrand, Vista University, Central University Technology, CGS, CSIR, ARC, SAAO, Natal Museum, Monash SA University, HMO, HSRC, NECSA, NHLS, MRC, NRF, Mintek, National Museum Bloemfontein, Sasol, DBSA.

The scientific fields considered are:

Agricultural sciences, Health Sciences, Biological sciences, Pharmaceutical Sciences, Chemical sciences, Technologies and applied sciences, Law, Engineering sciences, Physical sciences, Social Sciences, Medical sciences: Basic, Arts, Humanities, Earth and marine sciences, Mathematical sciences, Information and Computer science, Economic sciences, Medical sciences: Clinical.

### A.3 Robustness check

I further examine homophilous preferences of the groups looking yearly student-supervisor ties. As robustness check, I use a logistic regression approach. For consistency as above, I remove the departments without population variability and that have per period fewer than 10 supervisions. The presented logistic regression models predict the likelihood of same-type tie versus cross-type tie with the following structure:

$$\text{logit}(p_{ij}) = \alpha + \beta X_i + \gamma X_j + \delta X_{ij} \quad (1)$$

Where  $ij$  is the tie between student  $i$  and professor  $j$ . The dependent variable is equal to 1 if the student and the advisor are of the same race (gender) and zero otherwise.  $X_i$ ,  $X_j$  are agents' covariates and  $X_{ij}$  are the covariates of the tie. The main independent variables are the dummies student and professor race (gender), both equal to 1 for white (male) and zero otherwise. As additional control I include student and professor gender (race), the student enrollment year, the fraction of black (female) students and professors (computed in the 5 time periods considered), professor age, professor rating<sup>1</sup> (a control for professor "quality")<sup>2</sup>, and department dummies (institutional constraints of university and field). Looking at homophilous preferences of the different groups I interpret, for example, a positive coefficient of *studrace* means that white students more likely to form same-type ties compared to black students and thus have stronger homophilous preferences than black students.

Table 5.4 shows in models 1 and 2 (3 and 4) results using as dependent variable same-race (same-gender). I report estimation done in the whole sample and for the sub-sample future academics. Previous results are consistent. White (male) students form same-race (same-gender) ties more likely than white (male) ones; while black (female) advisors form same-race (same-gender) ties more likely than white (male) colleges. In contrast, in the sub-sample of future academics (model 2) white professors are more likely to form same-race ties than black professors. In table 1 and in the table 2 in the Appendix A I control for sub-population sizes and results are unchanged.

<sup>1</sup>Professor rating is the individual evaluation of the research quality of a professor and it covers the period 1983 – 2012. The NRF grades researchers following an examination of a candidate's research output. The process involves international referees who evaluate the CV and published papers of the professor. This process ends with a rating: one of 7 ordered categories. See Section 2.3.

<sup>2</sup>I have to remark that this variable reduces the sample. To account for this, I estimated the presented models also excluding the variable and I found the same results. The rating variable does not show a systematic effect.



**Table 1:** Results of logistic regressions of yearly student-supervisor racial tie data. Dependent variable same-race is 1 for same-race ties and zero for cross-race ties. Models 3 and 4 are for the sub-sample of future scholars. Faculties with fewer than 10 observations and without population variability are excluded. Where variable studrace (studgender) is a students' dummy 1 for white (male); race (gender) is a professors' dummy 1 for white (male). Age is professor's age; YearFirstRegistration is the students' registration year. Sb and Tb are proportions of black students and professors aggregated for each time period. Universities + Field are department dummies, and Rating is professors' NRF rating (7 rating categories).

	<i>Dependent variable:</i>			
	same-race ties			
	(1)	(2)	(3)	(4)
studrace	3.025*** (0.050)	3.027*** (0.050)	2.655*** (0.105)	2.659*** (0.105)
race	-0.834*** (0.050)	-0.837*** (0.050)	1.053*** (0.117)	1.052*** (0.117)
studgender	-0.092** (0.043)	-0.092** (0.043)	-0.024 (0.103)	-0.021 (0.103)
gender	0.197*** (0.045)	0.195*** (0.045)	0.193* (0.104)	0.193* (0.104)
age	0.001 (0.003)	0.002 (0.003)	-0.003 (0.006)	-0.003 (0.006)
YearFirstRegistration	-0.008 (0.015)	-0.016 (0.016)	0.022 (0.037)	0.008 (0.038)
Sb	-1.257 (1.838)	-0.213 (1.927)	-0.902 (11.861)	-5.404 (12.286)
Tb	1.526 (1.584)	-8.331 (5.564)	0.921 (8.285)	-16.390 (14.659)
Sb:Tb		12.410* (6.715)		27.080 (18.899)
Constant	14.864 (30.474)	32.915 (32.008)	-46.270 (72.927)	-14.099 (76.235)
Universities + Field	yes	yes	yes	yes
Rating	yes	yes	yes	yes
Observations	14,924	14,924	2,934	2,934
Log Likelihood	-7,351.759	-7,350.050	-1,411.695	-1,410.666
Akaike Inf. Crit.	14,833.520	14,832.100	2,939.390	2,939.332

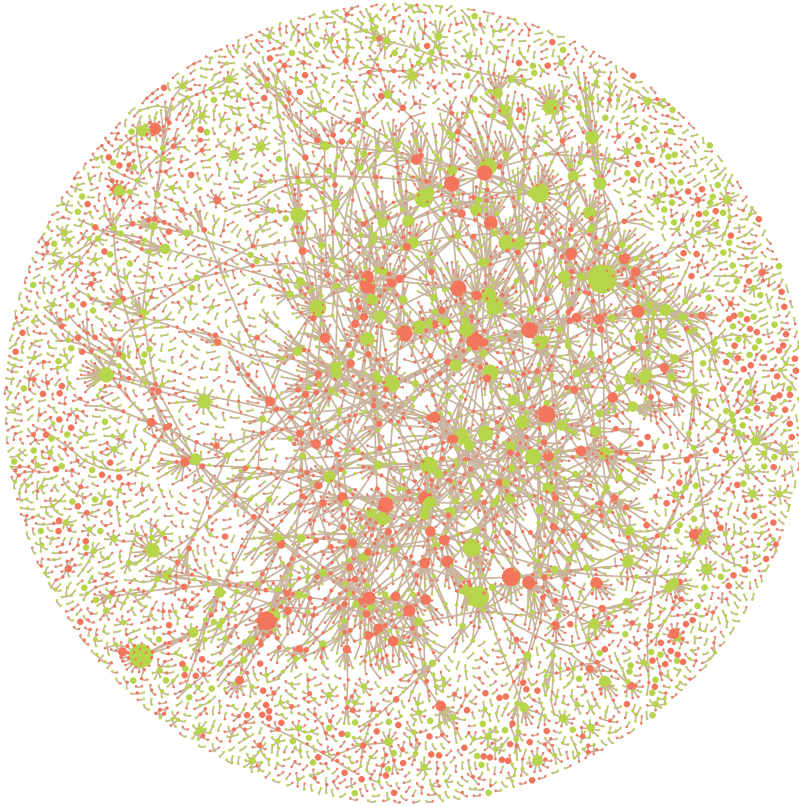
Note:

\*p<0.1; \*\*p<0.05; \*\*\*p<0.01

## A.4 Gender Analysis

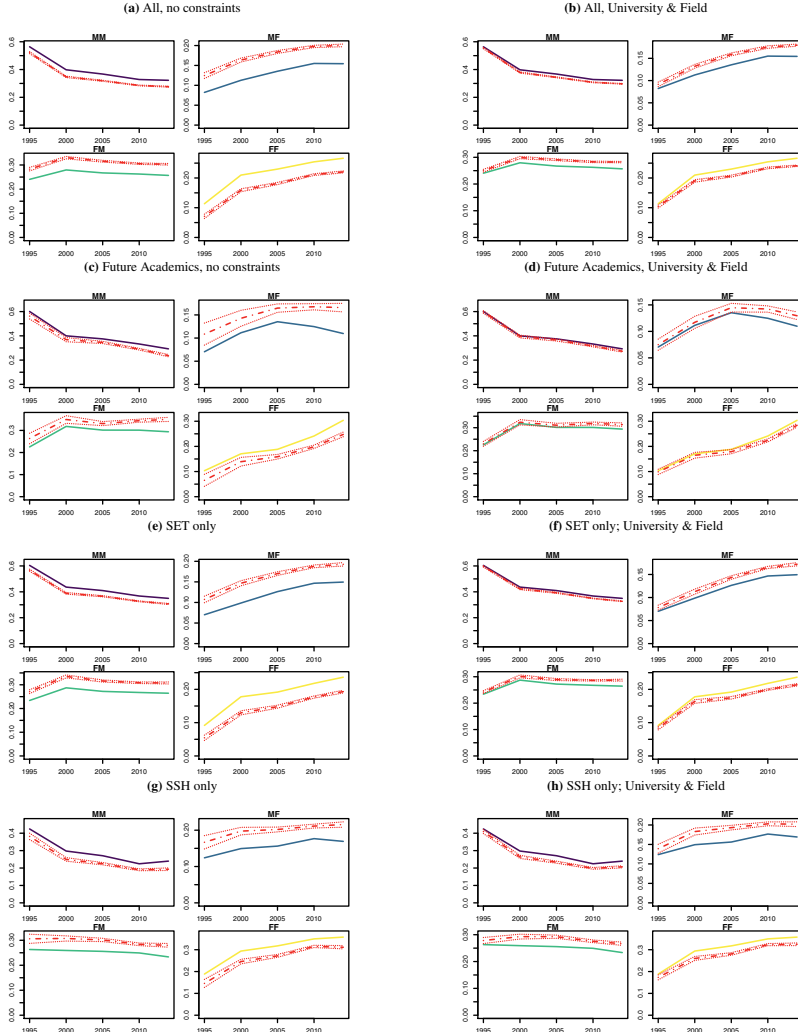
### A.4.1 Future Academics

**Figure 3:** Student-Supervisor network for Future Academics 1973-2014. The size of the nodes are proportional to node degree. The colour of the nodes represents gender: coral are female, green are male

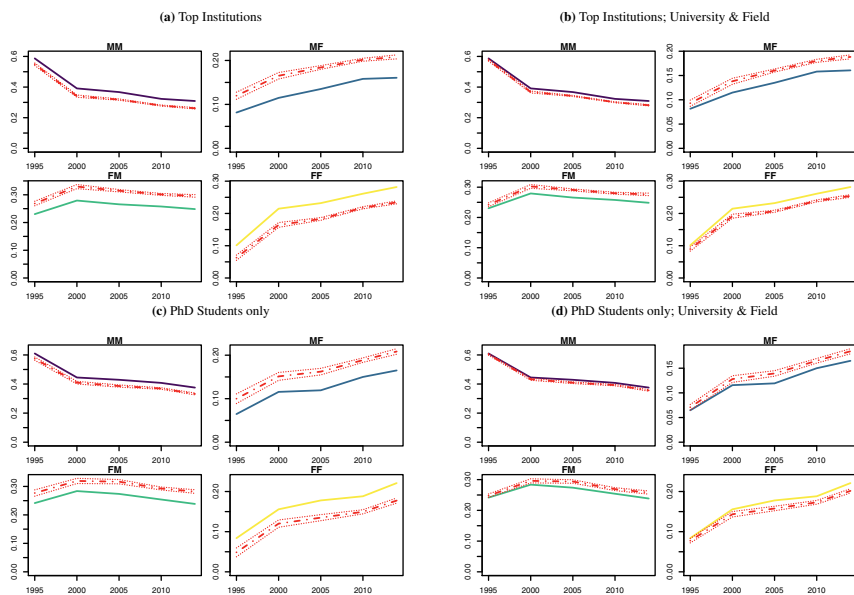


## A.4.2 Permutation Test

**Figure 4:** Permutation test for student supervisor data 1973-2014. The permutation is done keeping fixed the number of ties and the supervisor and permuting the students. Permutation without constraint (left) and with institutional constraints (right). Permutation is repeated 100 times for each of time period. The Observed proportions of gender student-supervisor relations (solid lines) are plotted versus the results of the permutation test (dashed lines) with respective two standard deviation from the mean.

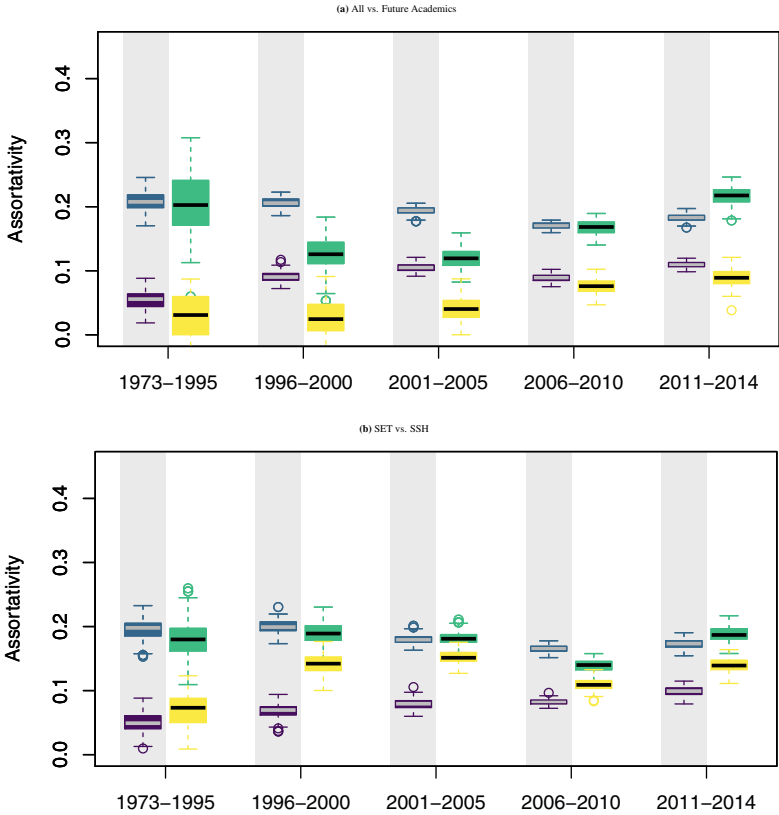


**Figure 5:** Permutation test for student supervisor data 1973-2014 for top institutions and PhD students. The permutation is done keeping fixed the number of ties and the supervisor and permuting the students. Permutation without constraint (left) and with institutional constraints (right). Permutation is repeated 100 times for each of time period. The Observed proportions of gender student-supervisor relations (solid lines) are plotted versus the results of the permutation test (dashed lines) with respective two standard deviation from the mean. As top institutions we include: Cape Town; Pretoria KwaZuluNatal; Stellenbosch; Rhodes; Witwatersrand; and Western Cape University;"

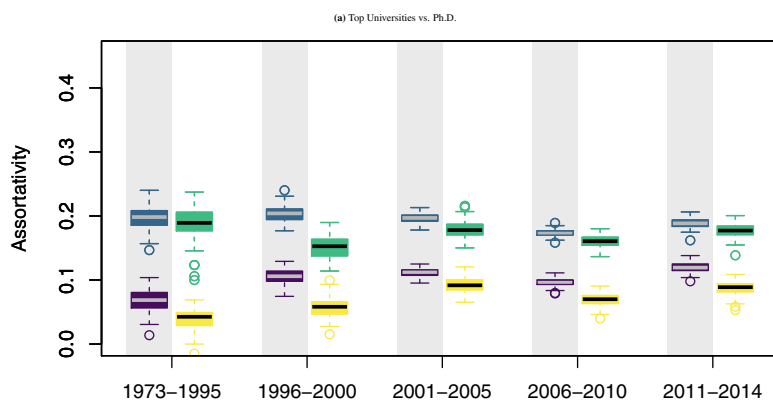


A.4.3 Choice and Induced Homophily

**Figure 6:** Induced and Choice Homophily comparison looking gender ties in student supervisor relations 1973-2014. We compare results of assortativity of different sub-samples: All vs. Future Academics (a), SET vs. SSH (b). Each panel contains 4 series; In Fig.(a): aggregated (upper series) versus department level (lower series) crossed with total sample (white background) versus future academics (grey background). And similarly for Fig. (b). Upper series are system-level permutations, including both induced and choice homophily; lower series are department-level permutations, excluding much of the induced homophily. Each box plot represents 100 permutations and associated assortativity calculation, as described in section 2.4.2, and equation 2.4.

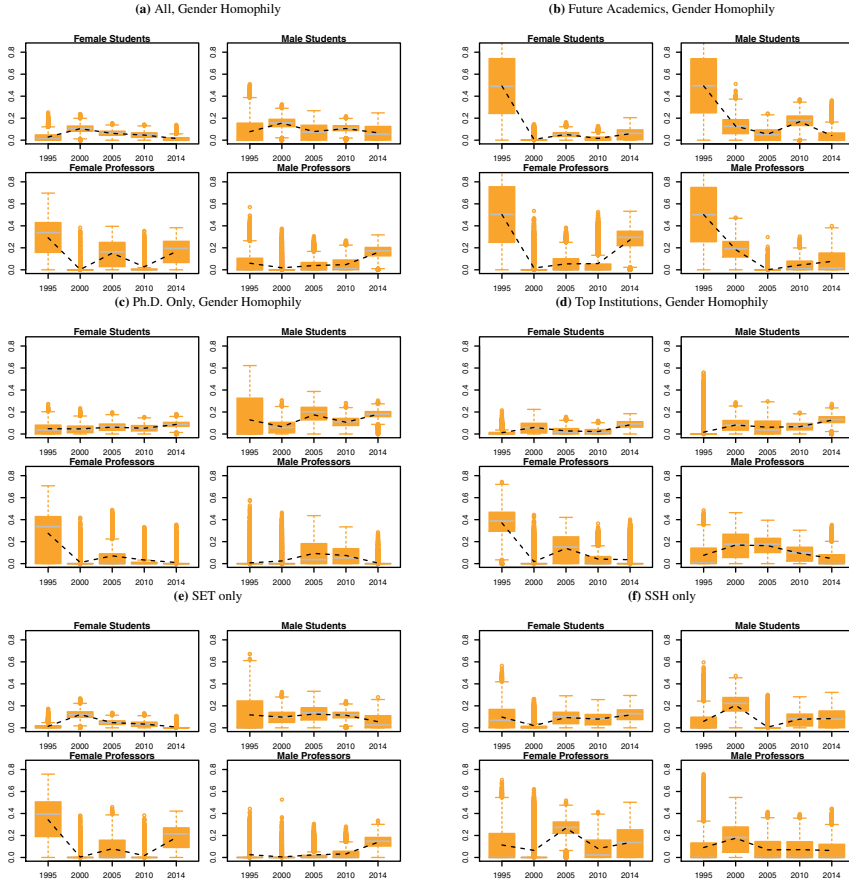


**Figure 7:** Induced and Choice Homophily comparison looking gender ties in student supervisor relations 1973-2014. We compare results of assortativity of two sub-samples: Top Universities vs. Ph.D. Sub-sample. The panel contains 4 series: aggregated (upper series) versus department level (lower series) crossed with Top Universities (white background) versus Ph.D. (grey background). Upper series are system-level permutations, including both induced and choice homophily; lower series are department-level permutations, excluding much of the induced homophily. Each box plot represents 100 permutations and associated assortativity calculation, as described in section 2.4.2, and equation 2.4.



### A.4.4 Choice Homophily of different groups

**Figure 8:** Gender Homophily of the different groups 1973-2014. The estimation is done minimising the average relative entropy of tie-types between observed and predicted from equations 2.6. We use data at department level under 10000 samples with replacement of department observations. Black dashed are averages. Faculties with less than 10 observations and without population variability are excluded.



**Table 2:** Results of logistic regressions of yearly student-supervisor gender tie data. Dependent variable same-gender is 1 for same-gender ties and zero for cross-race ties. Models 3 and 4 are for the sub-sample of future scholars. Faculties with fewer than 10 observations and without population variability are excluded. Where variables: studrace (studgender) is a students' dummy 1 for white (male); race (gender) is a professors' dummy 1 for white (male). Age is professor's age; YearFirstRegistration is the students' registration year. Sf and Tf are proportions of female students and professors aggregated for each time period. Universities + Field are department dummies, and Rating is professors' NRF rating (7 rating categories).

	<i>Dependent variable:</i>			
	same-gender ties			
	(1)	(2)	(3)	(4)
studrace	0.058** (0.026)	0.058** (0.026)	0.114* (0.065)	0.112* (0.065)
race	0.024 (0.031)	0.024 (0.031)	-0.141* (0.082)	-0.140* (0.082)
studgender	0.893*** (0.024)	0.893*** (0.024)	1.140*** (0.060)	1.141*** (0.060)
gender	-0.605*** (0.026)	-0.604*** (0.026)	-0.960*** (0.063)	-0.959*** (0.063)
age	-0.002 (0.002)	-0.002 (0.002)	-0.014*** (0.004)	-0.013*** (0.004)
YearFirstRegistration	-0.011 (0.007)	-0.016* (0.009)	0.006 (0.019)	-0.011 (0.022)
Sf	1.440 (6.722)	-40.396 (39.172)	-0.712 (15.845)	-138.266 (93.336)
Tf	0.636 (2.603)	-40.419 (37.963)	1.321 (6.234)	-131.649 (89.135)
Sf:Tf		86.907 (80.169)		282.468 (188.883)
Constant	20.846 (15.588)	51.648 (32.408)	-11.425 (39.345)	88.826 (77.716)
Universities + Field	yes	yes	yes	yes
Rating	yes	yes	yes	yes
Observations	33,310	33,310	6,026	6,026
Log Likelihood	-21,569.580	-21,568.990	-3,752.535	-3,751.416
Akaike Inf. Crit.	43,325.160	43,325.980	7,691.070	7,690.832

Note:

\*p<0.1; \*\*p<0.05; \*\*\*p<0.01





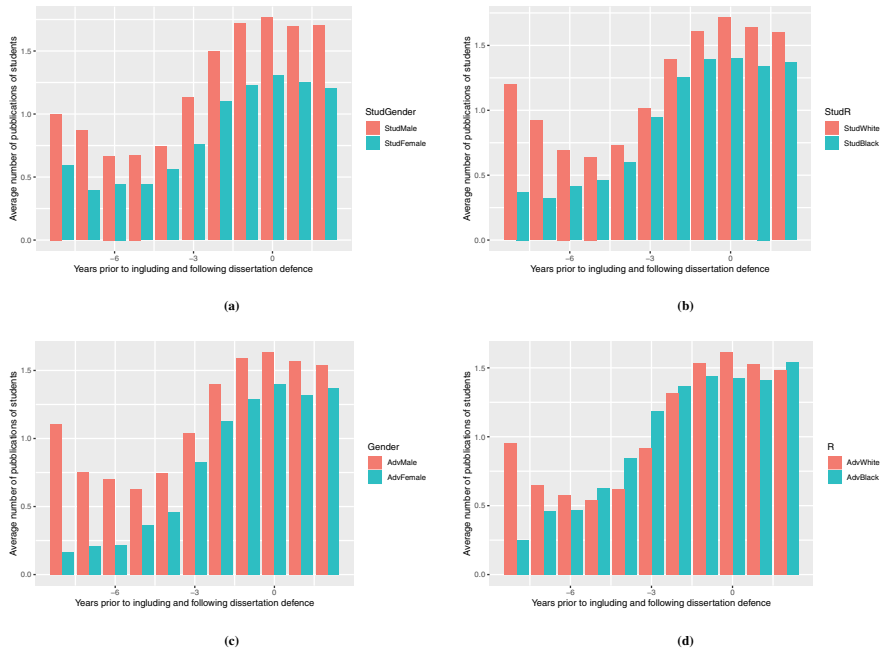
## B.5 Variables

### Independent variable:

- **moreAdv** is a dummy variable equal to 1 if the students has more then one supervisor. One third of the students have more, the maximum number of advisor per student is 3
- **logprofcumavgprod** is the log of 1+ the lagged cumulative average productivity of advisor. The average cumulative number of paper is computed since the year of the first record in the publication data to t-1 and divided by the number of years.
- **DummyStudPrevPub** is a dummy equal to 1 if the student has published before. Overall the 28% of male students has already publish before starting the Ph.D.; while for female student this percentage is 25%. This suggest that the gap in publication could be originated before starting the Ph.D.
- **timegrad** time to graduation

## B.6 Additional Statistics on the data

**Figure 9:** Average three years publications of students classified by gender(a) and race(b) of the student and gender(c) and race(d) of advisor. The average for the groups is calculated every year starting from 8 years before the thesis defence (d-8) until two years after the thesis defence (d+2). Where year of defence (d) is equal to zero.



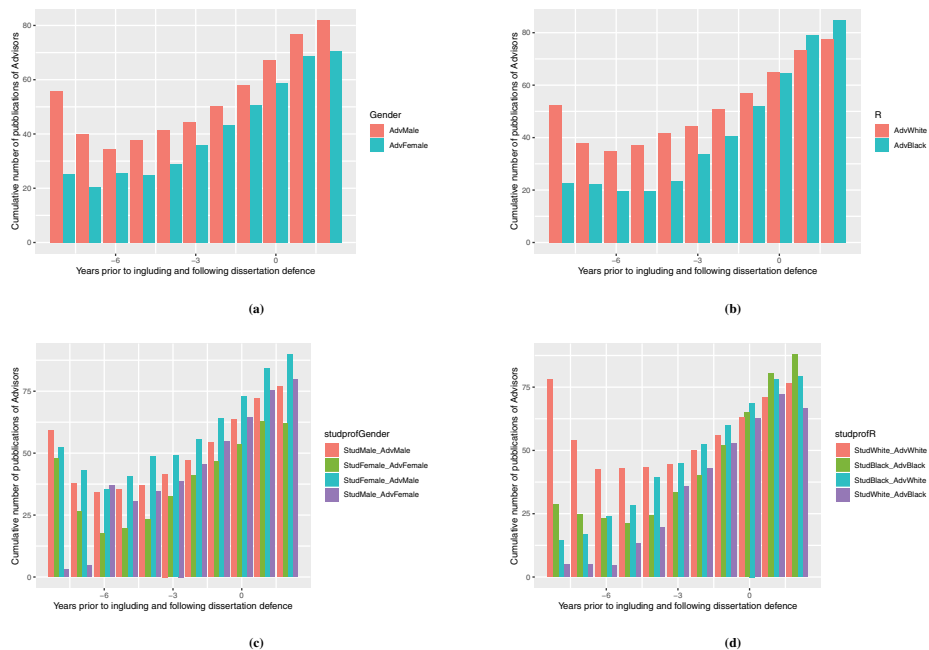
**Table 3:** Student 3 year average number of publications including and excluding zeros for white male, white female, black male, and black female. The logarithms are showed at the bottom of the table, excluding the zeros.

	White Male	White Female	F/M	Black Male	Black Female	F/M
Including zeros						
Mean	1.58	1.21	0.77	1.37	0.75	0.55
Median	0.33	0		0	0	
Std dev.	2.87	3.30		2.81	2.12	
Obs.	1636	1673		1840	900	
Excluding zeros						
Mean	3.16	2.86	0.91	2.89	1.82	0.63
Median	2.00	1.33	0.67	1.67	1.00	0.60
Std dev.	3.38	4.58		3.50	3.00	
Obs.	820	707		873	370	
In logarithms Excluding zeros						
Mean	1.17	1.04	-0.13	1.10	0.81	-0.29
Median	1.10	0.85	-0.25	0.98	0.69	-0.29
Std dev.	0.69	0.69		0.67	0.57	
Obs.	820	707		873	370	

**Table 4:** Advisor Logarithm of 1+ cumulative average productivity from first record to t-1. It refers to the variable called Logprofcumavgprod

	White Male	White Female	F-M	Black Male	Black Female	F-M
Including zeros						
Mean	1.29	1.14	-0.15	1.33	1.67	0.34
Median	1.36	1.17	-0.19	1.33	1.61	0.28
Std dev.	0.78	0.77		0.96	0.86	
Obs.	3311	1413		1120	205	
Excluding zeros						
Mean	1.53	1.44	-0.09	1.67	1.77	0.1
Median	1.52	1.38	-0.14	1.58	1.65	0.07
Std dev.	0.59	0.57		0.76	0.77	
Obs.	2775	1108		881	193	

**Figure 10:** Cumulative number of publications of advisor classified by gender(a) and race(b) and by couple of student and advisor gender(c) and race(d). The average for the groups is calculated every year starting from 8 years before the thesis defence (d-8) until two years after the thesis defence (d+2). Where year of defence (d) is equal to zero.



**Table 5:** Distribution of the study sample by year of thesis defence

	Student	Advisor
2001	1	1
2002	7	7
2003	23	23
2004	36	32
2005	58	54
2006	80	75
2007	76	69
2008	89	80
2009	90	77
2010	114	93
2011	117	106
2012	127	110
2013	91	79
2014	15	14

**Table 6:** Distribution of Ph.D. students and advisors by the discipline of the thesis.

	Student	Black Stud.	White Stud.	Female Stud.	Male Stud	Advisor	Black Adv.	White Adv.	Female Adv.	Male Adv.
Agricultural sciences	90	46	44	44	46	63	15	48	45	18
Biological sciences	278	96	182	154	124	142	19	123	92	50
Chemical sciences	117	66	51	73	44	49	16	33	38	11
Earth and marine sciences	67	28	39	35	32	44	6	38	36	8
Engineering sciences	69	26	43	51	18	51	10	41	46	5
Health Sciences	81	33	48	34	47	59	18	41	33	26
Information and Computer science	31	5	26	16	15	21	3	18	17	4
Mathematical sciences	24	14	10	18	6	20	6	14	19	1
Medical sciences: Basic	43	25	18	17	26	28	11	17	20	8
Medical sciences: Clinical	14	7	7	6	8	12	2	10	6	6
Pharmaceutical Sciences	15	15	0	10	5	3	2	1	3	0
Physical sciences	55	37	18	49	6	36	9	27	31	5
Technologies and applied sciences	40	18	22	24	16	24	4	20	18	6

**Table 7:** Average productivity for different sub-sample of the data and student advisor couple. Productivity is computed as  $\log(1 + \text{pub}_t)$  where  $\text{pub}_t$  is number of student publication between years  $t$  and  $t + 2$  inclusive.

		Advisor		
		Male	Female	Average
Student	Male	0.65	0.44	0.61
	Female	0.44	0.46	0.44
	Average	0.56	0.45	0.53

(a) Same-race pair, white student white advisor

		Advisor		
		White	Black	Average
Student	White	0.65	0.29	0.63
	Black	0.46	0.6	0.52
	Average	0.57	0.56	0.57

(b) Same-gender pair, male student male advisor

		Advisor		
		Male	Female	Average
Student	Male	0.6	0.54	0.59
	Female	0.35	0.55	0.37
	Average	0.51	0.54	0.52

(c) Same-race pair, black student black advisor

		Advisor		
		White	Black	Average
Student	White	0.44	0.09	0.42
	Black	0.19	0.55	0.23
	Average	0.38	0.26	0.37

(d) Same-gender pair, female student female advisor

		Advisor		
		Male	Female	Average
Student	Male	0.46	0.53	0.48
	Female	0.37	0.19	0.31
	Average	0.43	0.4	0.42

(e) Cross-race pair, black student white advisor

		Advisor		
		White	Black	Average
Student	White	0.44	0.57	0.45
	Black	0.37	0.35	0.36
	Average	0.43	0.39	0.42

(f) Cross-gender pair, female student male advisor

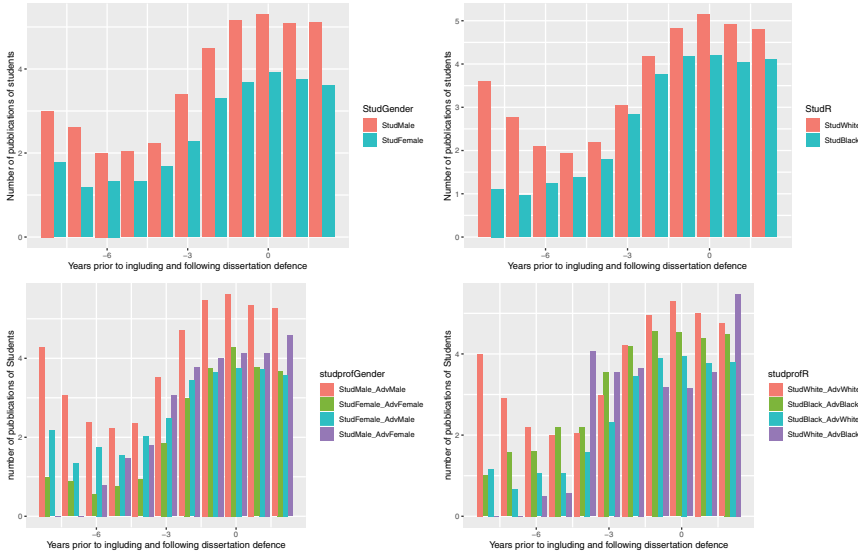
**Table 8:** Gender and Racial Assortativity Coefficient (Ass.) by universities and field. The assortativity coefficient is computed according to Newman (2003), while 95% confidence intervals are computed simulating 1000 times type-blind tie formation given supervision and population composition.

Univerisity	Assortativity coefficient by Universities				Ass. Race	sign	95% CI Null Model
	Ass. Gender	sign	95% CI Null Model				
CapePeninsulaUniversityOfTechnology;	-0.40		(-0.75 ; 0.65 )	0.00			(-1.00 ; 1.00 )
DurbanInstituteOfTechnology;	-0.39		(-0.57 ; 0.48 )	0.76	*		(-0.65 ; 0.53 )
NelsonMandelaMetropolitanUniversity;	0.30		(-0.41 ; 0.38 )	0.33	*		(-0.33 ; 0.33 )
NorthWestUniversity;	0.54	*	(-0.38 ; 0.36 )	0.47	*		(-0.47 ; 0.47 )
RhodesUniversity;	-0.11		(-0.28 ; 0.26 )	0.26	*		(-0.16 ; 0.17 )
TshwaneUniversityOfTechnology;	-0.45	*	(-0.45 ; 0.42 )	0.60	*		(-0.47 ; 0.47 )
UniversityOfCapeTown;	0.10		(-0.14 ; 0.13 )	0.19	*		(-0.13 ; 0.12 )
UniversityOfFortHare;	0.00		(0.00 ; 0.00 )	0.66	*		(-0.89 ; 0.66 )
UniversityOfJohannesburg;	0.17		(-0.39 ; 0.37 )	0.47	*		(-0.33 ; 0.29 )
UniversityOfKwaZuluNatal;	0.18		(-0.28 ; 0.28 )	0.46	*		(-0.30 ; 0.33 )
UniversityOfLimpopo;	0.00		(0.00 ; 0.00 )	1.00			NA
UniversityOfPretoria;	0.11		(-0.14 ; 0.15 )	0.20	*		(-0.15 ; 0.13 )
UniversityOfSouthAfrica;	1.00		(-1.00 ; 1.00 )	0.00			(0.00 ; 0.00 )
UniversityOfStellenbosch;	0.05		(-0.14 ; 0.14 )	0.20	*		(-0.13 ; 0.14 )
UniversityOfTheFreeState;	0.22	*	(-0.20 ; 0.22 )	0.05			(-0.11 ; 0.09 )
UniversityOfTheWesternCape;	-0.07		(-0.23 ; 0.25 )	0.02			(-0.24 ; 0.22 )
UniversityOfVenda;	0.33		(-1.00 ; 1.00 )	0.33			(-1.00 ; 1.00 )
UniversityOfWitwatersrand;	0.16		(-0.56 ; 0.40 )	0.00			(-0.31 ; 0.31 )
UniversityOfZululand;	1.00		NA	1.00			NA
VaalUniversityOfTechnology;	1.00		NA	1.00			NA

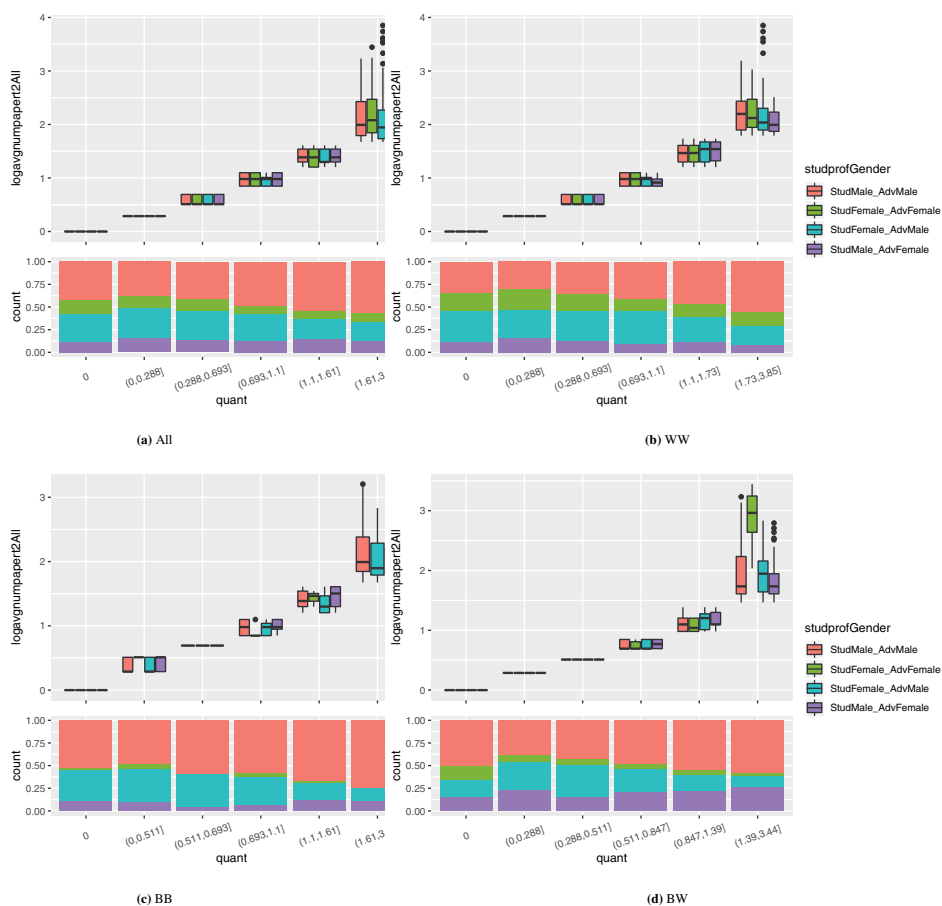
  

Field	Assortativity coefficient by Field				Ass. Race	sign	95% CI Null Model
	Ass. Gender	sign	95% CI Null Model				
Agricultural sciences	0.14		(-0.17 ; 0.19 )	0.52	*		(-0.19 ; 0.17 )
Biological sciences	0.06		(-0.12 ; 0.12 )	0.14	*		(-0.11 ; 0.11 )
Chemical sciences	-0.07		(-0.19 ; 0.17 )	0.46	*		(-0.16 ; 0.17 )
Earth and marine sciences	0.05		(-0.20 ; 0.17 )	0.19	*		(-0.19 ; 0.19 )
Engineering sciences	0.11		(-0.22 ; 0.20 )	0.52	*		(-0.23 ; 0.22 )
Health Sciences	0.15		(-0.21 ; 0.20 )	0.28	*		(-0.20 ; 0.20 )
Information and Computer science	0.08		(-0.32 ; 0.28 )	0.63	*		(-0.50 ; 0.50 )
Mathematical sciences	0.24	*	(-0.36 ; 0.24 )	0.14			(-0.33 ; 0.38 )
Medical sciences: Basic	0.02		(-0.24 ; 0.23 )	0.55	*		(-0.30 ; 0.28 )
Medical sciences: Clinical	0.29		(-0.57 ; 0.57 )	0.29			(-0.43 ; 0.43 )
Pharmaceutical Sciences	0.00		(0.00 ; 0.00 )	0.40			(-0.80 ; 0.60 )
Physical sciences	0.16		(-0.42 ; 0.42 )	0.26	*		(-0.18 ; 0.19 )
Technologies and applied sciences	0.06		(-0.28 ; 0.28 )	0.41	*		(-0.23 ; 0.25 )

**Figure 11:** Three years number of publications of students classified by student gender/race (top) and couple student-advisor gender/race of advisor (bottom). The average for the groups is calculated every year starting from 8 years before the thesis defence (d-8) until two years after the thesis defence (d+2).



**Figure 12:** Distribution of the dependent variable over the deciles of its distribution by student-advisor gender couple. The bottom stack-plots represent the relative proportions of the gender couple in the population for each decile. (a) Whole sample, (b) sub-sample of white students with white advisors, (c) sub-sample of black students with black advisors, and (c) sub-sample of black students with white advisors.





## B.7 Regression with gender and race interaction on the whole sample

**Table 9:** Pooled OLS Panel Regression with robust clustered standard error on the whole sample with interaction terms of gender and race. The dependent variable is log of 1+ average productivity in terms of number of paper between the period  $t$  and  $t+2$ . Models (1) for the Students comparison; Model (2) Advisor comparison; Model (3) Couples Comparison. Additional controls are moreAdv, logprofcumavgprod, DummyStudPrevPub, timegrad, field, and enrolment year.

	(1)	(2)	(3)
StudFemale	-0.105*		
	(0.0480)		
StudBlack	-0.00433		
	(0.0497)		
StudFemaleBlack	-0.0277		
	(0.0703)		
AdvFemale		-0.0291	
		(0.0452)	
AdvBlack		-0.000940	
		(0.0504)	
AdvFemaleBlack		-0.0787	
		(0.113)	
StudFemale_AdvFemale			-0.0855
			(0.0716)
StudFemale_AdvMale			-0.128*
			(0.0594)
StudMale_AdvFemale			-0.0508
			(0.0862)
StudBlack_AdvBlack			0.0154
			(0.0767)
StudBlack_AdvWhite			-0.0115
			(0.0665)
StudWhite_AdvBlack			-0.000968
			(0.129)
StudFemale_AdvFemale # StudBlack_AdvBlack			-0.386
			(0.213)
StudFemale_AdvFemale # StudBlack_AdvWhite			-0.0425
			(0.139)
StudFemale_AdvFemale # StudWhite_AdvBlack			-0.229
			(0.151)
StudFemale_AdvMale # StudBlack_AdvBlack			-0.0572
			(0.106)
StudFemale_AdvMale # StudBlack_AdvWhite			0.0234
			(0.0968)
StudFemale_AdvMale # StudWhite_AdvBlack			-0.0577
			(0.263)
StudMale_AdvFemale # StudBlack_AdvBlack			0.0102
			(0.189)
StudMale_AdvFemale # StudBlack_AdvWhite			-0.0446
			(0.120)
StudMale_AdvFemale # StudWhite_AdvBlack			-0.185
			(0.219)
Constant	0.516***	0.393**	0.409***
	(0.150)	(0.140)	(0.112)
N	6049	6049	6049
R <sup>2</sup>	0.285	0.281	0.287

Standard errors in parentheses

\*  $p < 0.05$

\*\*  $p < 0.01$

\*\*\*  $p < 0.001$

## B.8 Main OLS regressions

**Table 10:** Pooled OLS Panel Regression with robust clustered standard error. The dependent variable is log of 1+ average productivity in terms of number of paper between the period t and t+2. Models (1) for the Students comparison; Model (2) Advisor comparison; Model (3) Couples Comparison. Where Columns (a) On the sub-sample White Student White Professors; Columns (b) On the sub-sample Black Student Black Professors; Columns (c) On the sub-sample Black Student White Professors. Additional controls are field, and enrolment year.

	(1) ALL	(1)a WW	(1)b BB	(1)c BW	(2) ALL	(2)a WW	(2)b BB	(2)c BW	(3) ALL	(3)a WW	(3)b BB	(3)c BW
StudFemale	-0.113** (0.0357)	-0.116* (0.0497)	-0.217* (0.0847)	-0.0908 (0.0686)								
AdvFemale					-0.0393 (0.0416)	-0.0248 (0.0589)	0.0989 (0.151)	-0.0678 (0.0720)	-0.120* (0.0571)	-0.113 (0.0745)	-0.270 (0.239)	-0.123 (0.129)
StudFemale_AdvFemale												
StudFemale_AdvMale									-0.136** (0.0417)	-0.142* (0.0580)	-0.198* (0.0958)	-0.112 (0.0810)
StudMale_AdvFemale									-0.0733 (0.0568)	-0.0622 (0.0899)	0.0768 (0.172)	-0.0931 (0.0839)
moreAdv	0.0585 (0.0431)	0.133* (0.0621)	0.143 (0.108)	-0.0929 (0.0675)	0.0525 (0.0434)	0.134* (0.0629)	0.0934 (0.113)	-0.0974 (0.0678)	0.0583 (0.0430)	0.131* (0.0628)	0.143 (0.113)	-0.0938 (0.0684)
logprofcumavgprod	0.122*** (0.0286)	0.151*** (0.0369)	0.0544 (0.0651)	0.143* (0.0643)	0.122*** (0.0287)	0.153*** (0.0370)	0.0492 (0.0666)	0.140* (0.0643)	0.122*** (0.0287)	0.151*** (0.0370)	0.0528 (0.0664)	0.143* (0.0646)
DummyStudPrevPub	0.791*** (0.0443)	0.825*** (0.0608)	0.559*** (0.101)	0.772*** (0.0921)	0.796*** (0.0443)	0.834*** (0.0611)	0.579*** (0.101)	0.790*** (0.0933)	0.790*** (0.0442)	0.822*** (0.0608)	0.566*** (0.102)	0.788*** (0.0952)
timegrad	-0.0401*** (0.0121)	-0.0387* (0.0168)	-0.0511 (0.0340)	-0.0236 (0.0241)	-0.0401*** (0.0122)	-0.0383* (0.0167)	-0.0594 (0.0348)	-0.0236 (0.0244)	-0.0395*** (0.0121)	-0.0388* (0.0168)	-0.0525 (0.0343)	-0.0201 (0.0241)
Constant	0.520*** (0.120)	0.294* (0.144)	1.535** (0.493)	0.489* (0.204)	0.407*** (0.118)	0.148 (0.155)	1.190* (0.483)	0.463* (0.209)	0.421*** (0.108)	0.197 (0.132)	1.323** (0.465)	0.416* (0.187)
N	6049	3083	1099	1641	6049	3083	1099	1641	6049	3083	1099	1641
R <sup>2</sup>	0.284	0.354	0.260	0.263	0.280	0.349	0.244	0.260	0.285	0.355	0.260	0.264

Standard errors in parentheses

\*\* p<0.01

\*\*\* p<0.001"

## B.9 Poisson panel regressions Gender

### B.9.1 Students

**Table 11:** Poisson Panel Regression with robust clustered standard error. The dependent variable is number of papers between the period  $t$  and  $t+2$ . Column (1) On the sub-sample White Student White Professors; Column (2) On the sub-sample Black Student Black Professors; Column (3) On the sub-sample Black Student White Professors. Additional controls are field, and enrolment year.

	ALL	(1) WW	(2) BB	(3) BW
StudFemale	-0.408*** (0.124)	-0.485** (0.166)	-0.871** (0.324)	-0.547* (0.243)
moreAdv	0.0734 (0.140)	0.264 (0.218)	0.292 (0.294)	-0.317 (0.290)
logprofcumavgprod	0.427*** (0.0965)	0.468*** (0.135)	0.186 (0.201)	0.805** (0.264)
DummyStudPrevPub	1.655*** (0.117)	1.896*** (0.165)	1.448*** (0.278)	2.071*** (0.291)
timegrad	-0.148*** (0.0424)	-0.0920 (0.0559)	-0.201 (0.124)	0.0476 (0.0881)
Constant	1.415*** (0.398)	0.257 (0.530)	4.766*** (0.798)	0.447 (0.704)
/				
lnalpha	1.471*** (0.400)	1.375* (0.547)	1.199 (0.902)	1.442 (0.838)
N	6049	3083	1099	1641

Standard errors in parentheses

\*  $p < 0.05$

\*\*  $p < 0.01$

\*\*\*  $p < 0.001$

## B.9.2 Advisors

**Table 12:** Poisson Panel Regression with robust clustered standard error. The dependent variable is number of papers between the period t and t+2. Column (1) On the sub-sample White Student White Professors; Column (2) On the sub-sample Black Student Black Professors; Column (3) On the sub-sample Black Student White Professors. Additional controls are field, and enrolment year.

	ALL	(1) WW	(2) BB	(3) BW
AdvFemale	-0.00650 (0.155)	-0.0205 (0.187)	-0.0947 (0.382)	-0.302 (0.294)
moreAdv	0.0960 (0.142)	0.260 (0.221)	0.0887 (0.291)	-0.309 (0.302)
logprofcumavgprod	0.428*** (0.0966)	0.468*** (0.135)	0.187 (0.201)	0.802** (0.266)
DummyStudPrevPub	1.664*** (0.117)	1.932*** (0.168)	1.410*** (0.277)	2.116*** (0.309)
timegrad	-0.150*** (0.0419)	-0.0870 (0.0534)	-0.264* (0.129)	0.0753 (0.102)
Constant	0.919* (0.404)	-0.468 (0.519)	4.069*** (0.901)	0.0270 (0.717)
/				
lnalpha	1.481*** (0.398)	1.391* (0.548)	1.230 (0.897)	1.451 (0.833)
N	6049	3083	1099	1641
Standard errors in parentheses				
* p<0.05	** p<0.01		*** p<0.001"	

### B.9.3 Student and Advisor couple

**Table 13:** Poisson Panel Regression with robust clustered standard error. The dependent variable is number of papers between the period  $t$  and  $t+2$ . Column (1) On the sub-sample White Student White Professors; Column (2) On the sub-sample Black Student Black Professors; Column (3) On the sub-sample Black Student White Professors. Additional controls are field, and enrolment year.

	ALL	(1) WW	(2) BB	(3) BW
StudFemale_AdvFemale	-0.335 (0.209)	-0.420 (0.216)	-1.303** (0.505)	-0.992 (0.683)
StudFemale_AdvMale	-0.437** (0.143)	-0.504* (0.202)	-0.828* (0.345)	-0.581* (0.279)
StudMale_AdvFemale	0.0162 (0.200)	0.0635 (0.292)	-0.0460 (0.414)	-0.362 (0.313)
moreAdv	0.0755 (0.140)	0.262 (0.218)	0.267 (0.298)	-0.309 (0.291)
logprofcumavgprod	0.427*** (0.0965)	0.468*** (0.136)	0.187 (0.201)	0.801** (0.265)
DummyStudPrevPub	1.658*** (0.117)	1.900*** (0.167)	1.480*** (0.283)	2.135*** (0.305)
timegrad	-0.146*** (0.0421)	-0.0903 (0.0551)	-0.207 (0.127)	0.0915 (0.100)
Constant	0.997** (0.359)	-0.245 (0.480)	3.884*** (0.747)	-0.0736 (0.652)
/				
lnalpha	1.471*** (0.400)	1.375* (0.548)	1.198 (0.902)	1.437 (0.839)
N	6049	3083	1099	1641

Standard errors in parentheses

\*  $p < 0.05$

\*\*  $p < 0.01$

\*\*\*  $p < 0.001$

## B.10 Quantile Regressions Gender

**Table 14:** Quantile regression with clustered standard errors on the complete sample. The dependent variable is log of 1+ average productivity in terms of number of paper between the period t and t+2. Additional controls are field, enrolment year, and year.

	(1)25th	(1)50th	(1)70th	(1)75th	(1)80th	(1)85th	(1)90th	(1)95th	(1)99th
StudFemale_AdvFemale	-2.44e-15 (3.76e-15)	-3.18e-15 (5.17e-15)	-0.129* (0.0605)	-0.223** (0.0766)	-0.295** (0.0905)	-0.330*** (0.0969)	-0.285* (0.128)	-0.224 (0.129)	-0.0191 (0.108)
StudFemale_AdvMale	-2.79e-16 (3.22e-15)	5.04e-15 (2.63e-15)	-0.126* (0.0578)	-0.194** (0.0723)	-0.247** (0.0801)	-0.269** (0.0832)	-0.276** (0.0887)	-0.255* (0.117)	-0.195** (0.0754)
StudMale_AdvFemale	3.64e-16 (4.93e-15)	-5.11e-16 (3.95e-15)	-0.0713 (0.0596)	-0.127 (0.0803)	-0.180 (0.103)	-0.180 (0.112)	-0.166 (0.115)	-0.0843 (0.139)	-0.110 (0.0950)
moreAdv	-2.09e-15 (3.54e-15)	1.60e-15 (2.96e-15)	0.0767 (0.0568)	0.0574 (0.0626)	0.0726 (0.0818)	0.0830 (0.0848)	0.0669 (0.100)	0.105 (0.116)	0.00277 (0.0733)
logprofcumavgprod	6.71e-15 (3.84e-15)	-2.11e-15 (1.75e-15)	0.127*** (0.0383)	0.138*** (0.0352)	0.154*** (0.0398)	0.181*** (0.0476)	0.209*** (0.0512)	0.219*** (0.0528)	0.175** (0.0549)
DummyStudPrevPub	0.511*** (1.91e-14)	0.981*** (2.16e-14)	1.069*** (0.0921)	1.031*** (0.0789)	0.968*** (0.0740)	0.946*** (0.0880)	0.927*** (0.0890)	0.913*** (0.123)	0.752*** (0.0669)
timegrad	3.38e-15 (2.13e-15)	8.65e-16 (1.11e-15)	-0.0117 (0.0145)	-0.0311 (0.0176)	-0.0527** (0.0186)	-0.0790*** (0.0209)	-0.104*** (0.0261)	-0.145*** (0.0272)	-0.179*** (0.0162)
N	6012	6012	6012	6012	6012	6012	6012	6012	6012
R <sup>2</sup>	0.246	0.230	0.306	0.312	0.312	0.303	0.294	0.272	0.196

Standard errors in parentheses

\* p<0.05

\*\* p<0.01

\*\*\* p<0.001\*

**Table 15:** Quantile regression with clustered standard errors on the sub-sample of White Student and White Advisor. The dependent variable is log of 1+ average productivity in terms of number of paper between the period t and t+2. Additional controls are field, enrolment year, and year.

	(1)25th	(1)50th	(1)70th	(1)75th	(1)80th	(1)85th	(1)90th	(1)95th	(1)99th
StudFemale_AdvFemale	-2.45e-15 (2.01e-15)	6.67e-15 (0.00646)	-0.0715 (0.0921)	-0.160 (0.0998)	-0.213 (0.110)	-0.238 (0.147)	-0.271* (0.137)	-0.361* (0.175)	-0.409*** (0.100)
StudFemale_AdvMale	-1.61e-15 (1.56e-15)	3.67e-15 (0.00471)	-0.114 (0.0801)	-0.174 (0.0895)	-0.204* (0.0949)	-0.245 (0.139)	-0.250 (0.163)	-0.334* (0.143)	-0.467*** (0.0749)
StudMale_AdvFemale	1.24e-16 (2.41e-15)	2.99e-15 (0.00698)	0.00621 (0.109)	-0.0367 (0.116)	-0.0230 (0.159)	-0.0541 (0.176)	-0.173 (0.142)	-0.237 (0.170)	-0.386*** (0.0588)
moreAdv	4.20e-15 (2.54e-15)	1.03e-14 (0.00564)	0.0991 (0.0679)	0.112 (0.0885)	0.122 (0.0927)	0.129 (0.117)	0.0814 (0.169)	0.00501 (0.133)	-0.0306 (0.0710)
logprofcumavgprod	-3.49e-16 (8.33e-16)	-5.02e-16 (0.00286)	0.108* (0.0425)	0.145** (0.0516)	0.162** (0.0583)	0.181* (0.0794)	0.236* (0.0951)	0.192** (0.0664)	0.184*** (0.0216)
DummyStudPrevPub	0.511*** (4.34e-15)	1.099*** (0.0328)	1.192*** (0.114)	1.122*** (0.123)	1.067*** (0.118)	0.958*** (0.116)	0.924*** (0.171)	0.758*** (0.125)	0.495*** (0.0711)
timegrad	-2.31e-16 (4.39e-16)	3.72e-17 (0.00169)	-0.00440 (0.0162)	-0.0214 (0.0162)	-0.0396* (0.0191)	-0.0625** (0.0211)	-0.0920** (0.0337)	-0.135*** (0.0361)	-0.176*** (0.0118)
N	3058	3058	3058	3058	3058	3058	3058	3058	3058
R <sup>2</sup>	0.312	0.298	0.364	0.376	0.380	0.373	0.357	0.317	0.247

Standard errors in parentheses

\* p<0.05

\*\* p<0.01

\*\*\* p<0.001\*

**Table 16:** Quantile regression with clustered standard errors on the subsample of Black Student and Black Advisor. The dependent variable is log of 1+ average productivity in terms of number of paper between the period t and t+2. Additional controls are field, enrolment year, and year.

	(2)25th	(2)50th	(2)70th	(2)75th	(2)80th	(2)85th	(2)90th	(2)95th	(2)99th
StudFemale_AdvFemale	1.09e-14 (0.0840)	-0.0751 (0.144)	-0.396 (0.272)	-0.300 (0.228)	-0.438* (0.218)	-0.532* (0.224)	-0.659** (0.218)	-0.919*** (0.160)	-0.865*** (0.0788)
StudFemale_AdvMale	1.99e-15 (0.0364)	-0.0245 (0.0381)	-0.254 (0.189)	-0.291* (0.128)	-0.351* (0.144)	-0.342 (0.176)	-0.376** (0.133)	-0.440* (0.174)	-0.310*** (0.0426)
StudMale_AdvFemale	3.74e-15 (0.0561)	0.00603 (0.0586)	-0.0638 (0.200)	-0.0319 (0.210)	-0.159 (0.187)	-0.182 (0.160)	-0.145 (0.150)	-0.0549 (0.202)	-0.355*** (0.0365)
moreAdv	6.48e-17 (0.0474)	0.0172 (0.0443)	0.104 (0.146)	0.134 (0.134)	0.0726 (0.148)	0.0446 (0.170)	0.0712 (0.111)	-0.0222 (0.108)	-0.0622** (0.0234)
logprofcumavgprod	-2.52e-15 (0.0191)	0.0379 (0.0370)	0.127 (0.0797)	0.118 (0.0742)	0.133 (0.0836)	0.104 (0.0696)	0.0796 (0.0617)	0.103* (0.0450)	0.277*** (0.0176)
DummyStudPrevPub	0.490*** (0.0882)	0.801*** (0.126)	0.692** (0.220)	0.571*** (0.155)	0.572** (0.174)	0.654*** (0.198)	0.697*** (0.144)	0.655*** (0.103)	0.629*** (0.0353)
timegrad	-8.63e-16 (0.0116)	0.0158 (0.0236)	-0.0179 (0.0858)	-0.0596 (0.0644)	-0.0897 (0.0522)	-0.108* (0.0546)	-0.121* (0.0503)	-0.178*** (0.0456)	-0.185*** (0.00787)
N	1091	1091	1091	1091	1091	1091	1091	1091	1091
R <sup>2</sup>	0.166	0.177	0.273	0.268	0.240	0.244	0.234	0.215	0.187

Standard errors in parentheses

\* p<0.05

\*\* p<0.01

\*\*\* p<0.001

**Table 17:** Quantile regression with clustered standard errors on the subsample of Black Student and White Advisor. The dependent variable is log of 1+ average productivity in terms of number of paper between the period t and t+2. Additional controls are field, enrolment year, and year.

	(3)25th	(3)50th	(3)70th	(3)75th	(3)80th	(3)85th	(3)90th	(3)95th	(3)99th
StudFemale_AdvFemale	1.82e-15 (1.36e-15)	3.80e-15 (3.97e-15)	-0.167 (0.155)	-0.221 (0.195)	-0.349 (0.205)	-0.355 (0.215)	-0.445 (0.228)	-0.371* (0.149)	-0.433*** (0.0484)
StudFemale_AdvMale	3.03e-16 (6.21e-16)	7.79e-16 (3.07e-15)	-0.100 (0.160)	-0.176 (0.198)	-0.271 (0.139)	-0.338* (0.155)	-0.365* (0.149)	-0.299* (0.119)	-0.157*** (0.0329)
StudMale_AdvFemale	-2.95e-16 (1.09e-15)	-2.00e-17 (4.79e-15)	-0.0283 (0.161)	-0.0869 (0.205)	-0.0843 (0.141)	-0.118 (0.187)	-0.130 (0.151)	-0.153 (0.101)	-0.0521 (0.0405)
moreAdv	-3.22e-16 (7.26e-16)	-1.25e-15 (2.80e-15)	-0.0839 (0.114)	-0.135 (0.121)	-0.165 (0.103)	-0.190 (0.0971)	-0.271* (0.124)	-0.179 (0.105)	-0.0589 (0.0393)
logprofcumavgprod	5.65e-16 (5.41e-16)	2.69e-16 (1.91e-15)	0.139 (0.109)	0.186 (0.103)	0.191* (0.0813)	0.163* (0.0806)	0.111 (0.0936)	0.0983 (0.0544)	0.123*** (0.0147)
DummyStudPrevPub	0.288 (.)	0.847*** (1.90e-13)	0.924*** (0.254)	0.931*** (0.273)	0.927*** (0.207)	0.984*** (0.255)	1.038*** (0.212)	1.030*** (0.115)	0.900*** (0.0446)
timegrad	2.01e-16 (1.81e-16)	7.76e-16 (1.38e-15)	0.0109 (0.0560)	0.000221 (0.0520)	-0.00841 (0.0492)	-0.0159 (0.0546)	-0.0600 (0.0539)	-0.105** (0.0323)	-0.129*** (0.00932)
N	1637	1637	1637	1637	1637	1637	1637	1637	1637
R <sup>2</sup>	0.152	0.186	0.281	0.287	0.275	0.274	0.243	0.178	0.147

Standard errors in parentheses

\* p<0.05

\*\* p<0.01

\*\*\* p<0.001

## B.11 Race Results

**Table 18:** Pooled OLS Panel Regression with robust clustered standard error. The dependent variable is log of 1+ average productivity in terms of number of paper between the period t and t+2. Models (1) for the Students comparison; Model (2) Advisor comparison; Model (3) Couples Comparison. Where Columns (a) On the sub-sample Male Student Male Professors; Columns (b) On the sub-sample Female Student Female Professors; Columns (c) On the sub-sample Male Student Male Professors. Additional controls are field, and enrolment year.

	(1) ALL	(1a) MM	(1b) FF	(1c) FM	(2) ALL	(2a) MM	(2b) FF	(2c) FM	(3) ALL	(3a) MM	(3b) FF	(3c) FM
StudBlack	0.00637 (0.0366)	0.0175 (0.0581)	-0.0567 (0.102)	-0.0128 (0.0623)								
AdvBlack					-0.00875 (0.0455)	0.0439 (0.0657)	-0.371* (0.156)	-0.0533 (0.0766)				
StudBlack_AdvBlack									0.00567 (0.0526)	0.0467 (0.0784)	-0.482 (0.289)	-0.0571 (0.0772)
StudBlack_AdvWhite									-0.00308 (0.0416)	0.00510 (0.0687)	-0.00275 (0.111)	0.0314 (0.0731)
StudWhite_AdvBlack									-0.0787 (0.0932)	0.0433 (0.124)	-0.320* (0.153)	0.0288 (0.246)
moreAdv	0.0515 (0.0436)	0.129 (0.0744)	-0.00657 (0.0803)	0.0402 (0.0798)	0.0515 (0.0434)	0.133 (0.0747)	-0.0442 (0.0790)	0.0442 (0.0806)	0.0518 (0.0437)	0.133 (0.0749)	-0.0462 (0.0822)	0.0419 (0.0821)
logprofcumavggrad	0.122*** (0.0287)	0.182*** (0.0502)	0.0714 (0.0455)	0.0649 (0.0530)	0.122*** (0.0287)	0.183*** (0.0503)	0.0749 (0.0453)	0.0637 (0.0529)	0.122*** (0.0288)	0.183*** (0.0503)	0.0735 (0.0456)	0.0638 (0.0532)
DummyStudPrevPub	0.798*** (0.0445)	0.853*** (0.0682)	0.835*** (0.114)	0.743*** (0.0890)	0.797*** (0.0444)	0.853*** (0.0671)	0.863*** (0.110)	0.741*** (0.0909)	0.796*** (0.0449)	0.854*** (0.0692)	0.876*** (0.115)	0.741*** (0.0889)
timegrad	-0.0407*** (0.0121)	-0.0640*** (0.0189)	-0.0470 (0.0263)	-0.00598 (0.0192)	-0.0405*** (0.0121)	-0.0637*** (0.0188)	-0.0392 (0.0270)	-0.00599 (0.0192)	-0.0400*** (0.0121)	-0.0637*** (0.0191)	-0.0388 (0.0274)	-0.00641 (0.0192)
Constant	0.348** (0.111)	0.414* (0.185)	0.315 (0.220)	0.250 (0.195)	0.367*** (0.109)	0.379* (0.182)	0.572** (0.207)	0.298 (0.206)	0.351*** (0.104)	0.422* (0.167)	0.213 (0.197)	0.244 (0.180)
N	6049	2683	825	1748	6049	2683	825	1748	6049	2683	825	1748
R <sup>2</sup>	0.279	0.317	0.417	0.298	0.279	0.317	0.433	0.297	0.279	0.317	0.435	0.299

Standard errors in parentheses

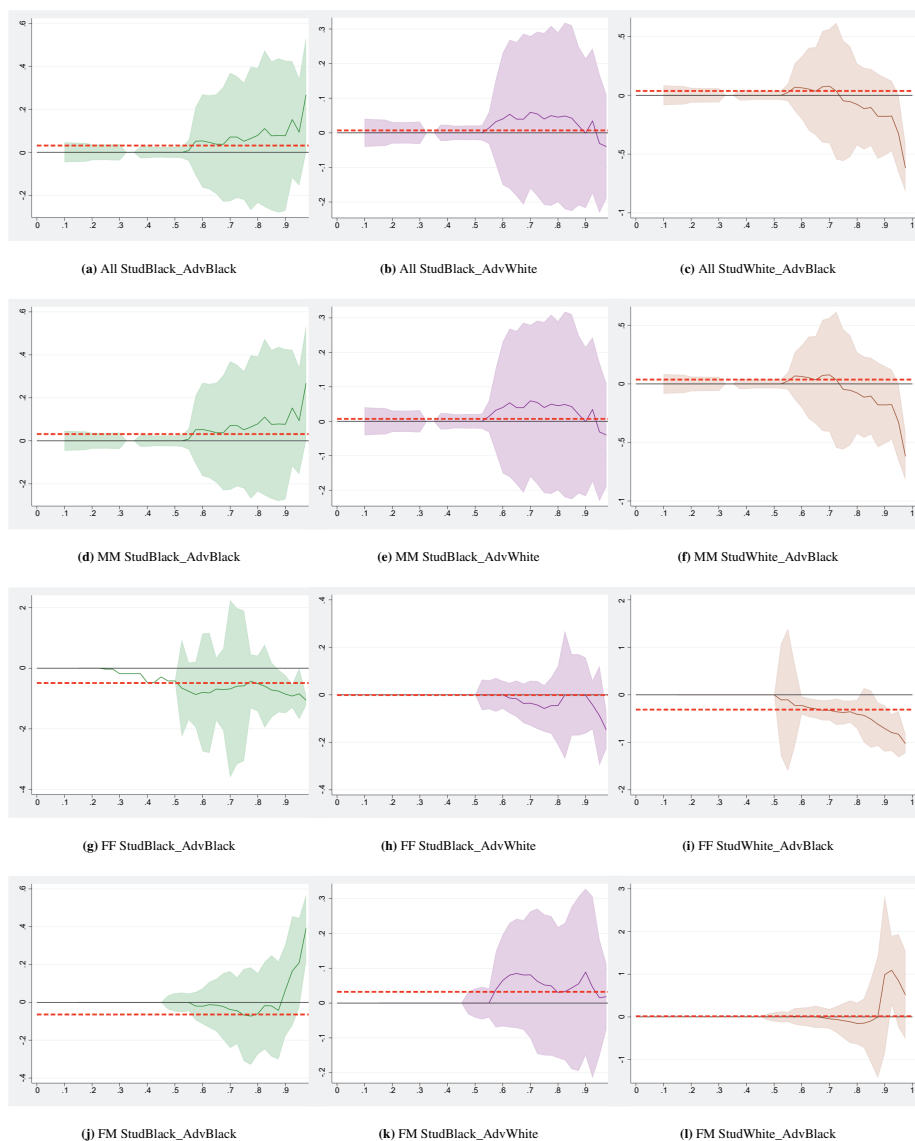
\* p&lt;0.05

\*\* p&lt;0.01

\*\*\* p&lt;0.001\*



**Figure 13:** Quantile Regression with clustered standard errors. Results for group comparison where the baseline group is White Student with White Advisor. Quantile regressions are done for each 2.5 percentile. Full lines are zero, dotted lines are panel OLS estimation of Models (3) in table 18. Additional controls are: discipline, enrolment year, year, time to graduation, whether the student had published previously, whether the student have more than one advisor, the log of average publications of the advisor lagged one year.



**Table 19:** Quantile regression with robust clustered standard errors where a, b, and c are respectively 75th, 80th, and 99th. The dependent variable is log of 1+ average productivity in terms of number of paper between the period t and t+2. (1) is for the sub-sample White Student White Professors; (2) is the sub-sample Black Student Black Professors; and (3) is the sub-sample Black Student White Professors. Additional controls are field, enrolment year, and year.

	(1)a 75th MM	(1)b 80th MM	(1)c 99th MM	(2)a 75th FF	(2)b 80th FF	(2)c 99th FF	(3)a 75th FM	(3)b 80th FM	(3)c 99th FM
StudBlack_AdvBlack	0.0513 (0.142)	0.0789 (0.161)	0.267** (0.0962)	-0.584 (1.171)	-0.508 (0.468)	-1.262*** (0.0638)	-0.0672 (0.125)	-0.0604 (0.109)	0.383*** (0.0376)
StudBlack_AdvWhite	0.0399 (0.128)	0.0452 (0.129)	-0.0989 (0.0531)	-0.0573 (0.0503)	-0.0451 (0.0848)	-0.136*** (0.0158)	0.0522 (0.103)	0.0302 (0.0949)	0.0649** (0.0232)
StudWhite_AdvBlack	0.0471 (0.260)	0.0774 (0.179)	0.581*** (0.126)	0.373** (0.125)	0.410** (0.155)	1.050*** (0.0666)	0.0868 (0.207)	0.152 (0.245)	-0.689*** (0.121)
moreAdv	0.0790 (0.137)	0.0955 (0.149)	0.293*** (0.0704)	-0.0256 (0.0646)	-0.0120 (0.0887)	0.131*** (0.0151)	0.107 (0.181)	0.203 (0.166)	0.0132 (0.0280)
logprofcumavgprod	0.299*** (0.0616)	0.327*** (0.0603)	0.237*** (0.0307)	0.0417 (0.0458)	4.04e-16 (0.0624)	0.0495* (0.0228)	0.0676 (0.0708)	0.0779 (0.0603)	0.126*** (0.0182)
DummyStudPrevPub	0.943*** (0.133)	0.909*** (0.147)	0.920*** (0.0871)	1.263*** (0.321)	1.299*** (0.232)	1.347*** (0.0555)	1.034*** (0.185)	1.080*** (0.128)	1.019*** (0.0300)
timegrad	-0.0852** (0.0319)	-0.115*** (0.0324)	-0.204*** (0.0382)	0.00337 (0.0291)	-9.92e-16 (0.0307)	0.00537 (0.00750)	0.0150 (0.0265)	0.00432 (0.0267)	-0.0122 (0.00984)
N	2668	2668	2668	811	811	811	1746	1746	1746
R <sup>2</sup>	0.338	0.326	0.241	0.389	0.368	0.262	0.316	0.307	0.224

Standard errors in parentheses

\* p<0.05

\*\* p<0.01

\*\*\* p<0.001"

## B.12 Poisson panel regressions race

### B.12.1 Students

**Table 20:** Poisson Panel Regression with robust clustered standard error. The dependent variable is number of papers between the period t and t+2. Column (1) On the sub-sample Male Student Male Professors; Column (2) On the sub-sample Female Student Female Professors; Column (3) On the sub-sample Female Student Male Professors. Additional controls are field, and enrolment year.

	ALL	(1) MM	(2) FF	(3) FM
StudBlack	0.203 (0.129)	0.157 (0.183)	-0.689 (0.622)	0.522 (0.268)
moreAdv	0.102 (0.143)	0.501* (0.221)	0.268 (0.448)	0.112 (0.287)
logprofcumavgprod	0.428*** (0.0965)	0.399** (0.146)	0.701* (0.301)	0.425** (0.138)
DummyStudPrevPub	1.700*** (0.122)	1.776*** (0.175)	3.188*** (0.537)	2.333*** (0.272)
timegrad	-0.155*** (0.0417)	-0.156** (0.0575)	0.0229 (0.135)	0.0250 (0.0816)
Constant	0.557 (0.409)	0.753 (0.683)	-2.456 (1.796)	-1.247 (0.753)
/				
lnalpha	1.479*** (0.397)	1.364* (0.580)	1.291 (1.137)	1.356 (0.753)
N	6049	2683	825	1748

Standard errors in parentheses

\* p<0.05

\*\* p<0.01

\*\*\* p<0.001"

### B.12.2 Advisors

**Table 21:** Poisson Panel Regression with robust clustered standard error. The dependent variable is number of papers between the period t and t+2. Column (1) On the sub-sample Male Student Male Professors; Column (2) On the sub-sample Female Student Female Professors; Column (3) On the sub-sample Female Student Male Professors. Additional controls are field, and enrolment year.

	ALL	(1) MM	(2) FF	(3) FM
AdvBlack	0.152 (0.151)	0.210 (0.199)	-3.972*** (0.892)	0.122 (0.311)
moreAdv	0.0969 (0.141)	0.498* (0.215)	0.0101 (0.444)	0.0643 (0.277)
logprofcumavgprod	0.428*** (0.0965)	0.399** (0.146)	0.715* (0.299)	0.426** (0.137)
DummyStudPrevPub	1.675*** (0.118)	1.774*** (0.168)	3.746*** (0.675)	2.221*** (0.256)
timegrad	-0.148*** (0.0420)	-0.148* (0.0605)	0.0574 (0.139)	0.0354 (0.0841)
Constant	0.673 (0.390)	0.700 (0.656)	0.765 (1.543)	-0.603 (0.656)
/				
lnalpha	1.480*** (0.397)	1.364* (0.578)	1.166 (1.164)	1.366 (0.757)
N	6049	2683	825	1748

Standard errors in parentheses

\* p<0.05

\*\* p<0.01

\*\*\* p<0.001 "

### B.12.3 Student and Advisor couple

**Table 22:** Poisson Panel Regression with robust clustered standard error . The dependent variable is number of papers between the period t and t+2. Column (1) On the sub-sample Male Student Male Professors; Column (2) On the sub-sample Female Student Female Professors; Column (3) On the sub-sample Female Student Male Professors. Additional controls are field, and enrolment year.

	ALL	(1) MM	(2) FF	(3) FM
StudBlack_AdvBlack	0.320 (0.181)	0.343 (0.244)	-3.422** (1.229)	0.400 (0.391)
StudBlack_AdvWhite	0.0997 (0.144)	0.0437 (0.198)	-0.211 (0.695)	0.666* (0.311)
StudWhite_AdvBlack	-0.400 (0.317)	-0.377 (0.489)	-4.190*** (0.838)	0.262 (0.568)
moreAdv	0.112 (0.142)	0.544* (0.224)	0.0651 (0.471)	0.0913 (0.296)
logprofcumavgprod	0.428*** (0.0965)	0.399** (0.146)	0.712* (0.300)	0.424** (0.138)
DummyStudPrevPub	1.719*** (0.123)	1.777*** (0.180)	3.635*** (0.688)	2.316*** (0.276)
timegrad	-0.148*** (0.0416)	-0.150* (0.0596)	0.0603 (0.138)	0.0201 (0.0866)
Constant	0.686 (0.351)	0.775 (0.590)	-3.266* (1.497)	-0.713 (0.649)
/				
lnalpha	1.476*** (0.397)	1.361* (0.579)	1.163 (1.174)	1.353 (0.754)
N	6049	2683	825	1748

Standard errors in parentheses

\* p<0.05

\*\* p<0.01

\*\*\* p<0.001"

## B.13 Quantile Regressions Race

**Table 23:** Quantile regression with clustered standard errors on the whole sample. The dependent variable is log of 1+ average productivity in terms of number of paper between the period t and t+2. Additional controls are field, enrolment year, and year.

	(1)25th	(1)50th	(1)70th	(1)75th	(1)80th	(1)85th	(1)90th	(1)95th	(1)99th
StudBlack_AdvBlack	-9.38e-16 (4.27e-15)	-4.55e-18 (2.78e-15)	-0.0105 (0.0636)	0.00431 (0.106)	0.0443 (0.105)	0.0723 (0.145)	0.139 (0.115)	0.170 (0.127)	0.206 (0.120)
StudBlack_AdvWhite	6.27e-16 (3.19e-15)	1.77e-15 (2.30e-15)	0.00828 (0.0505)	0.0277 (0.0679)	0.0357 (0.0769)	0.0382 (0.0922)	0.0167 (0.0713)	-0.0148 (0.0914)	0.00657 (0.170)
StudWhite_AdvBlack	9.22e-16 (8.10e-15)	8.26e-15* (3.99e-15)	-0.114 (0.0612)	-0.124 (0.0951)	-0.112 (0.137)	-0.103 (0.142)	-0.112 (0.159)	-0.0827 (0.315)	-0.0101 (0.187)
moreAdv	-2.13e-15 (3.86e-15)	1.37e-15 (2.59e-15)	0.0583 (0.0598)	0.0529 (0.0864)	0.0642 (0.0993)	0.0754 (0.115)	0.0580 (0.0749)	0.0384 (0.115)	0.0146 (0.101)
logprofcumavgprod	6.99e-15 (3.72e-15)	-7.50e-16 (1.28e-15)	0.119*** (0.0336)	0.152*** (0.0422)	0.168*** (0.0475)	0.197*** (0.0533)	0.224*** (0.0452)	0.234*** (0.0582)	0.204* (0.0815)
DummyStudPrevPub	0.511*** (2.25e-14)	0.981*** (1.63e-14)	1.069*** (0.0816)	1.037*** (0.0959)	0.989*** (0.0908)	0.944*** (0.108)	0.962*** (0.0978)	0.906*** (0.118)	0.798*** (0.0674)
timegrad	3.47e-15 (2.18e-15)	1.13e-15 (8.56e-16)	-0.00852 (0.0134)	-0.0300 (0.0215)	-0.0553* (0.0269)	-0.0861*** (0.0222)	-0.115*** (0.0180)	-0.149*** (0.0197)	-0.190*** (0.0179)
N	6012	6012	6012	6012	6012	6012	6012	6012	6012
R <sup>2</sup>	0.246	0.230	0.299	0.306	0.308	0.299	0.287	0.260	0.202

Standard errors in parentheses

\* p<0.05

\*\* p<0.01

\*\*\* p<0.001\*

**Table 24:** Quantile regression with clustered standard errors on the sub-sample of Male Student and Male Advisor. The dependent variable is log of 1+ average productivity in terms of number of paper between the period t and t+2. Additional controls are field, enrolment year, and year.

	(1)25th	(1)50th	(1)70th	(1)75th	(1)80th	(1)85th	(1)90th	(1)95th	(1)99th
StudBlack_AdvBlack	-5.30e-15 (0.0183)	2.65e-15 (0.0126)	0.0711 (0.152)	0.0513 (0.139)	0.0789 (0.160)	0.0764 (0.176)	0.0771 (0.178)	0.0936 (0.127)	0.267** (0.0962)
StudBlack_AdvWhite	-4.30e-15 (0.0155)	6.80e-15 (0.0105)	0.0594 (0.112)	0.0399 (0.127)	0.0452 (0.125)	0.0423 (0.137)	-0.00101 (0.110)	-0.0307 (0.102)	-0.0989 (0.0531)
StudWhite_AdvBlack	-8.47e-15 (0.0308)	-6.90e-15 (0.0189)	0.0776 (0.248)	-0.0471 (0.260)	-0.0774 (0.177)	-0.104 (0.166)	-0.179 (0.169)	-0.327 (0.168)	-0.581*** (0.126)
moreAdv	-1.47e-15 (0.0175)	3.29e-15 (0.0124)	0.139 (0.138)	0.0790 (0.137)	0.0955 (0.141)	0.130 (0.196)	0.179 (0.174)	0.191 (0.122)	0.293*** (0.0704)
logprofcumavgprod	6.32e-15 (0.00848)	-1.70e-17 (0.00595)	0.277*** (0.0555)	0.299*** (0.0610)	0.327*** (0.0586)	0.316*** (0.0694)	0.306*** (0.0532)	0.265*** (0.0586)	0.237*** (0.0307)
DummyStudPrevPub	0.511*** (0.0427)	1.099*** (0.0600)	1.048*** (0.127)	0.943*** (0.131)	0.909*** (0.147)	0.926*** (0.183)	0.949*** (0.156)	0.924*** (0.137)	0.920*** (0.0871)
timegrad	3.13e-15 (0.00507)	4.75e-16 (0.00401)	-0.0555 (0.0295)	-0.0852** (0.0314)	-0.115*** (0.0300)	-0.131*** (0.0279)	-0.166*** (0.0301)	-0.194*** (0.0231)	-0.204*** (0.0382)
N	2668	2668	2668	2668	2668	2668	2668	2668	2668
R <sup>2</sup>	0.254	0.240	0.338	0.357	0.326	0.319	0.312	0.294	0.241

Standard errors in parentheses

\* p<0.05

\*\* p<0.01

\*\*\* p<0.001\*

**Table 25:** Quantile regression with clustered standard errors on the sub-sample of Female Student and Female Advisor. The dependent variable is log of 1+ average productivity in terms of number of paper between the period t and t+2. Additional controls are field, enrolment year, and year.

	(2)25th	(2)50th	(2)70th	(2)75th	(2)80th	(2)85th	(2)90th	(2)95th	(2)99th
StudBlack_AdvBlack	-0.0323 (.)	-0.421 (.)	-0.684 (1.487)	-0.584 (1.309)	-0.508 (0.468)	-0.721 (0.459)	-0.855*** (0.312)	-0.845* (0.426)	-1.262*** (0.0638)
StudBlack_AdvWhite	-3.07e-16 (1.48e-15)	2.74e-15 (3.69e-15)	-0.0352 (0.0481)	-0.0573 (0.0531)	-0.0451 (0.0848)	2.74e-16 (0.0869)	-0.00283 (0.0815)	-0.0874 (0.107)	-0.136*** (0.0158)
StudWhite_AdvBlack	1.42e-15 (2.17e-15)	2.03e-16 (1.59e-15)	-0.324** (0.109)	-0.373*** (0.131)	-0.410*** (0.155)	-0.511 (0.306)	-0.709*** (0.249)	-0.828*** (0.248)	-1.050*** (0.0666)
moreAdv	1.00e-16 (1.68e-15)	-3.68e-16 (2.04e-15)	-0.0270 (0.0699)	-0.0256 (0.0602)	-0.0120 (0.0887)	1.61e-15 (0.0752)	0.00822 (0.0905)	0.0747 (0.0833)	0.131*** (0.0151)
logprofcumavgrad	-1.02e-15 (1.74e-15)	-2.02e-16 (1.23e-15)	0.0349 (0.0450)	0.0417 (0.0444)	7.42e-16 (0.0624)	8.25e-16 (0.0596)	-0.0134 (0.0670)	-0.00524 (0.129)	0.0495* (0.0228)
DummyStudPrevPub	0.288*** (1.97e-14)	0.981*** (4.86e-14)	1.314*** (0.246)	1.263*** (0.369)	1.299*** (0.232)	1.345*** (0.218)	1.337*** (0.143)	1.264*** (0.209)	1.347*** (0.0555)
timegrad	-6.85e-16 (7.53e-16)	6.01e-18 (5.08e-16)	0.0110 (0.0156)	0.00337 (0.0291)	-8.64e-16 (0.0307)	-1.25e-15 (0.0293)	-0.00539 (0.0246)	0.00153 (0.0377)	0.00537 (0.00750)
N	811	811	811	811	811	811	811	811	811
R <sup>2</sup>	0.280	0.303	0.384	0.386	0.368	0.368	0.286	0.275	0.262

Standard errors in parentheses

\* p&lt;0.05

\*\* p&lt;0.01

\*\*\* p&lt;0.001

**Table 26:** Quantile regression with clustered standard errors on the sub-sample of Female Student and Male Advisor. The dependent variable is log of 1+ average productivity in terms of number of paper between the period t and t+2. Additional controls are field, enrolment year, and year.

	(3)25th	(3)50th	(3)70th	(3)75th	(3)80th	(3)85th	(3)90th	(3)95th	(3)99th
StudBlack_AdvBlack	-1.41e-16 (2.80e-15)	2.35e-16 (0.0245)	-0.0382 (0.103)	-0.0672 (0.125)	-0.0604 (0.109)	-0.0188 (0.137)	0.0648 (0.124)	0.209 (0.121)	0.383*** (0.0376)
StudBlack_AdvWhite	7.19e-16 (1.97e-15)	1.98e-16 (0.0212)	0.0811 (0.0928)	0.0522 (0.103)	0.0302 (0.0949)	0.0438 (0.119)	0.0890 (0.122)	0.0155 (0.0856)	0.0649** (0.0232)
StudWhite_AdvBlack	1.12e-15 (4.16e-15)	2.58e-15 (0.0543)	-0.0449 (0.117)	-0.0868 (0.206)	-0.152 (0.245)	-0.0958 (0.497)	0.990 (0.941)	0.844 (0.555)	0.689*** (0.121)
moreAdv	3.86e-16 (2.02e-15)	-2.26e-16 (0.0230)	0.0492 (0.103)	0.107 (0.180)	0.203 (0.166)	0.215 (0.237)	0.138 (0.170)	0.133 (0.143)	0.0132 (0.0280)
logprofcumavgrad	-2.79e-15 (1.92e-15)	-2.48e-16 (0.0139)	0.0682 (0.0634)	0.0676 (0.0710)	0.0779 (0.0603)	0.106 (0.0746)	0.168 (0.0859)	0.0919 (0.0728)	0.126*** (0.0182)
DummyStudPrevPub	0.288*** (6.06e-15)	0.693*** (0.0688)	0.959*** (0.160)	1.034*** (0.185)	1.080*** (0.128)	1.051*** (0.145)	1.021*** (0.122)	1.118*** (0.103)	1.019*** (0.0300)
timegrad	-8.47e-16 (6.98e-16)	-6.17e-16 (0.00645)	0.0156 (0.0209)	0.0150 (0.0266)	0.00432 (0.0267)	-0.0211 (0.0389)	-0.0343 (0.0390)	-0.0188 (0.0309)	-0.0122 (0.00984)
N	1746	1746	1746	1746	1746	1746	1746	1746	1746
R <sup>2</sup>	0.225	0.242	0.311	0.316	0.307	0.299	0.271	0.249	0.224

Standard errors in parentheses

\* p&lt;0.05

\*\* p&lt;0.01

\*\*\* p&lt;0.001

**Table 28:** Pooled Panel Regression with robust clustered standard error. The dependent variable is log of 1+ average productivity in terms of number of paper between the period t and t+2. Column (1) On the sub-sample White Student White Professors; Column (2) On the sub-sample Black Student Black Professors; Column (3) On the sub-sample Black Student White Professors. Additional controls are field, and enrolment year.

	(1)	(2)	(3)
AdvFemale	-0.119 (0.0714)	0.137 (0.146)	-0.0473 (0.0834)
Constant	0.431** (0.159)	1.198** (0.402)	0.571** (0.206)
N	3083	1099	1641
R <sup>2</sup>	0.0830	0.119	0.0764
Standard errors in parentheses			
* p<0.05                      ** p<0.01              *** p<0.001"			

## B.14 Without controls

**Table 27:** Pooled Panel Regression with robust clustered standard error. The dependent variable is log of 1+ average productivity in terms of number of paper between the period t and t+2. Column (1) On the sub-sample White Student White Professors; Column (2) On the sub-sample Black Student Black Professors; Column (3) On the sub-sample Black Student White Professors. Additional controls are field, and enrolment year.

	(1)	(2)	(3)
StudFemale	-0.195** (0.0613)	-0.236** (0.0845)	-0.169* (0.0813)
Constant	0.559*** (0.143)	1.659*** (0.393)	0.743*** (0.189)
N	3083	1099	1641
R <sup>2</sup>	0.0949	0.138	0.0893
Standard errors in parentheses			
* p<0.05                      ** p<0.01              *** p<0.001"			



**Table 29:** Pooled Panel Regression with robust clustered standard error. The dependent variable is log of 1+ average productivity in terms of number of paper between the period t and t+2. Column (1) On the sub-sample White Student White Professors; Column (2) On the sub-sample Black Student Black Professors; Column (3) On the sub-sample Black Student White Professors. Additional controls are field, and enrolment year.

	(1)	(2)	(3)
StudFemale_AdvFemale	-0.252** (0.0941)	-0.0632 (0.230)	-0.291* (0.137)
StudFemale_AdvMale	-0.238*** (0.0716)	-0.240** (0.0923)	-0.101 (0.0957)
StudMale_AdvFemale	-0.198 (0.107)	0.0319 (0.173)	0.0327 (0.0989)
Constant	0.426*** (0.125)	1.411*** (0.368)	0.585*** (0.167)
N	3083	1099	1641
R <sup>2</sup>	0.0997	0.139	0.0941

Standard errors in parentheses

\* p<0.05

\*\* p<0.01

\*\*\* p<0.001"

**Table 30:** Pooled Panel Regression with robust clustered standard error. The dependent variable is log of 1+ average productivity in terms of number of paper between the period t and t+2. Column (1) On the sub-sample Male Student Male Professors; Column (2) On the sub-sample Female Student Female Professors; Column (3) On the sub-sample Female Student Male Professors. Additional controls are field, and enrolment year.

	(1)	(2)	(3)
StudBlack	-0.137 (0.0702)	-0.146 (0.111)	-0.0791 (0.0770)
Constant	0.859*** (0.198)	0.337 (0.225)	0.599** (0.192)
N	2683	825	1748
R <sup>2</sup>	0.0544	0.183	0.0884

Standard errors in parentheses

\* p<0.05

\*\* p<0.01

\*\*\* p<0.001"

**Table 31:** Pooled Panel Regression with robust clustered standard error. The dependent variable is log of 1+ average productivity in terms of number of paper between the period t and t+2. Column (1) On the sub-sample Male Student Male Professors; Column (2) On the sub-sample Female Student Female Professors; Column (3) On the sub-sample Female Student Male Professors. Additional controls are field, and enrolment year.

	(1)	(2)	(3)
AdvBlack	-0.0603 (0.0792)	-0.252 (0.163)	-0.109 (0.0943)
Constant	0.732*** (0.188)	0.383 (0.207)	0.622** (0.203)
N	2683	825	1748
R <sup>2</sup>	0.0491	0.184	0.0861

Standard errors in parentheses

\* p<0.05

\*\* p<0.01

\*\*\* p<0.001"

**Table 32:** Pooled Panel Regression with robust clustered standard error. The dependent variable is log of 1+ average productivity in terms of number of paper between the period t and t+2. Column (1) On the sub-sample Male Student Male Professors; Column (2) On the sub-sample Female Student Female Professors; Column (3) On the sub-sample Female Student Male Professors. Additional controls are field, and enrolment year.

	(1)	(2)	(3)
StudBlack_AdvBlack	-0.107 (0.0933)	-0.0329 (0.262)	-0.141 (0.0949)
StudBlack_AdvWhite	-0.197* (0.0812)	-0.202 (0.120)	-0.0204 (0.0878)
StudWhite_AdvBlack	-0.361* (0.155)	-0.393* (0.161)	0.0174 (0.320)
Constant	0.710*** (0.171)	0.175 (0.166)	0.526** (0.164)
N	2683	825	1748
R <sup>2</sup>	0.0634	0.200	0.0890

Standard errors in parentheses

\* p<0.05

\*\* p<0.01

\*\*\* p<0.001"



## C.15 University reform in 2004

The South African university system saw a major reform in 2004. The reform merged and split university departments in the spirit of a geographical rationalization and racial integration. In the analysis I use post-merger names mostly because the data are more complete. More precisely, it is possible to make an accurate translation from pre- to post-merger names, but not from post- to pre-merger names, so by using pre-merger names I would lose a significant number of observations. Moreover, the use of post-merger names represents a value added of the work. It is a way to produce the prestige ranking of the South African universities that can be compared to the actual system. From the point of view of the analysis I note the following. The University of Johannesburg came into existence as the result of a merger between Rand Afrikaans University, Technikon Witwatersrand, and Vista University, where the latter two have almost no PhDs (3 in total) in the period. So using University of Johannesburg instead of its disaggregation pre-merger would not make much difference. Similarly, Nelson Mandela Metropolitan University was created by the merger of Port Elizabeth Technikon, University of Port Elizabeth, and Vista University where the sample is dominated by PhDs from Port Elizabeth. NorthWest University is a merger of University of the North West and Potchefstroom University, and the latter dominates PhD production, particularly if I restrict attention to SET where 32 PhDs are from Potchefstroom versus 6 from the University of the North West. The only possible problem could arise for the case of University of KwaZulu Natal which is the merger of University of Durban West Ville and University of Natal. Though, restricting to SET, Natal dominates with 32 PhDs versus 6 in Durban West Ville.

As a robustness check I redo the prestige ranking in SET with those observations for which I have full data using pre-merger names. Table 33 shows the results which are quite consistent to those of the full sample. I must also be cognizant of the fact that in the period there was to some extent a language divide which appears mitigated using post-merger names. Indeed, looking separately at English and Afrikaans language universities in Table 33 I can observe a more informative pattern. English universities have the same rankings here as in the main analysis. The exception being KZN: University of Natal is ranked 4th in the analysis using old names whereas KZN was ranked second among English universities in the main analysis. Afrikaans universities have almost the same ranking here as they do in the main analysis; and if UNW (NWO and Potchefstroom) are excluded.

I have also repeated the matched pairs analysis with pre-merger names, and notwithstanding the reduced sample size, results are in line with the main findings.

### Prestige Ranking with pre-merger names

University	Prestige Ranking	Language
NorthWestUniversity	1.873418	Afr
RhodesUniversity	2.721519	Eng
UniversityOfWitwatersrand	4.417722	Eng
UniversityOfCapeTown	4.506329	Eng
UniversityOfTheOrangeFreeState	4.658228	Afr
UniversityOfNatal	5.78481	Eng
UniversityOfPretoria	7.088608	Afr
PotchefstroomUniversity	7.721519	Afr
UniversityOfDurbanWestVille	8.35443	Eng
RandAfrikaansUniversity	10.658228	Afr
UniversityOfTheFreeState	11.139241	Afr
UniversityOfStellenbosch	12.101266	Afr
UniversityOfSouthAfrica	12.367089	Afr/Eng
MedicalUniversityOfSouthAfrica	13.139241	Afr/Eng
UniversityOfPortElizabeth	14.683544	Afr/Eng
UniversityOfTheWesternCape	14.78481	Eng

**Table 33:** Prestige Ranking for SET 1970-2004 using pre-merger names. The prestige ranking is in ascending order from the highest prestige which correspond to one. The number is the average of the orders with the maximum scores of the 10,000 repetitions. The algorithm is run on the adjacency matrix of pre-merger university names of the hiring network. Universities with fewer than 5 PhDs are excluded.

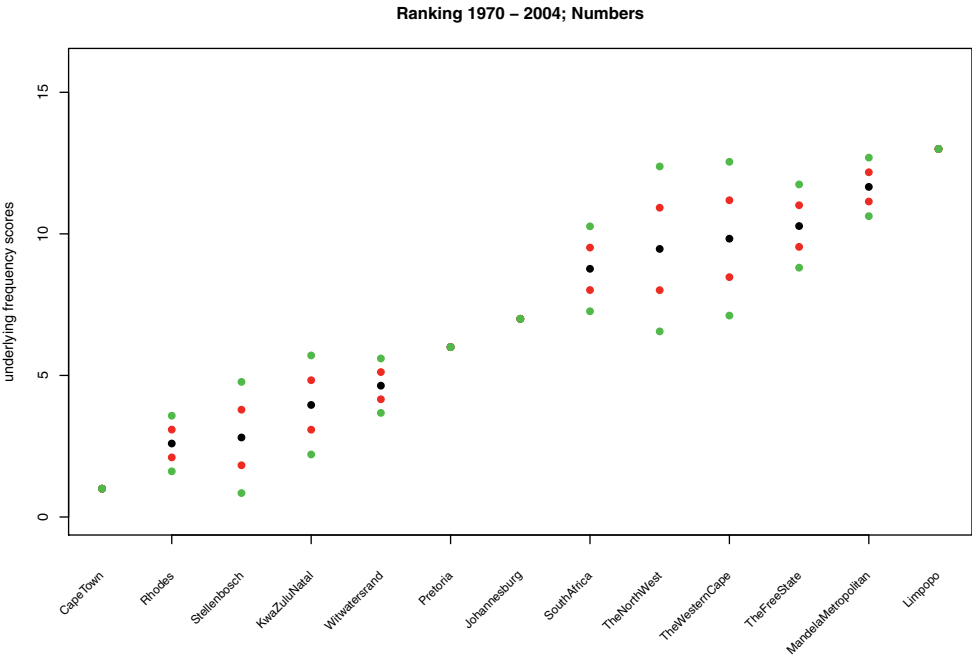
## C.16 Faculty hiring network

### Indegree and Outdegree

	SET		SSH	
	Indegree	Outdegree	Indegree	Outdegree
UniversityOfCapeTown	8	14	4	11
NelsonMandelaMetropolitanUniversity	8	9	6	2
UniversityOfWitwatersrand	9	11	4	6
UniversityOfPretoria	10	14	5	8
UniversityOfJohannesburg	8	10	5	4
UniversityOfTheFreeState	7	7	3	2
UniversityOfSouthAfrica	8	8	8	9
UniversityOfStellenbosch	10	15	5	14
UniversityOfKwaZuluNatal	7	13	8	8
UniversityOfTheNorthWest	8	10	4	6
UniversityOfLimpopo	11	3	7	0
RhodesUniversity	5	9	5	5
UniversityOfTheWesternCape	7	5	2	5
UniversityOfFortHare	7	1	2	0
WalterSisuluUniversity	3	0	5	1
UniversityOfVenda	3	1	2	0
CentralUniversityOfTechnology	1	1	1	1
TshwaneUniversityOfTechnology	6	0	4	1
VaalUniversityOfTechnology	2	1	2	0
MonashSAUniversity	0	1	0	0
DurbanInstituteOfTechnology	2	0	1	0
CapePeninsulaUniversityOfTechnology	3	0	0	0

**Table 34:** Indegree and Outdegree SET hiring network for the years 1970-2004. Network statistics are computed without considering self-loops

C.17 Prestige Ranking aggregating SET and SSH

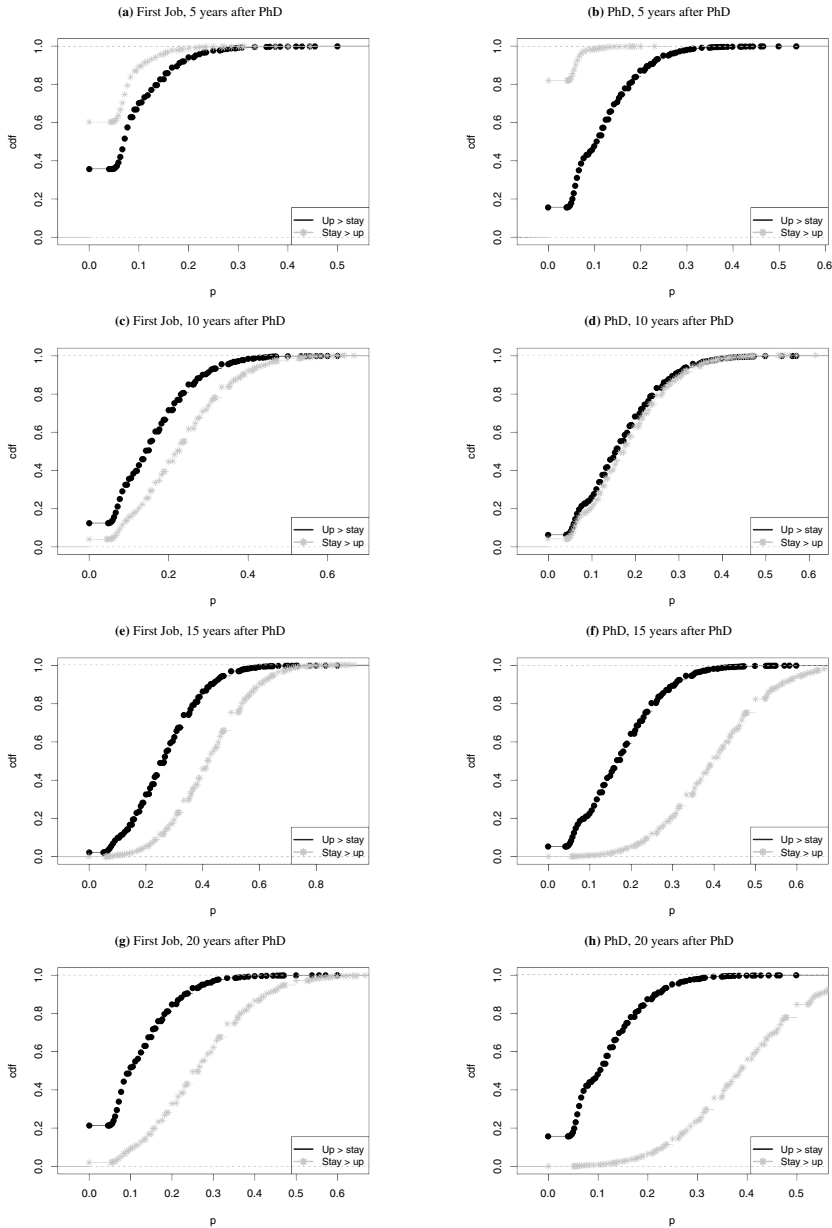


**Figure 14:** Prestige Ranking 1970-2004 without distinction of fields. The frequency scores are in ordered left to right from the highest prestige which corresponds to one. The Black dots represent the average placement of each university in the maintained orderings, red dots and green dots are respectively one and two standard deviation from the average. The algorithm runs on the adjacency matrix of the hiring network. Universities with fewer than 5 PhDs are excluded.

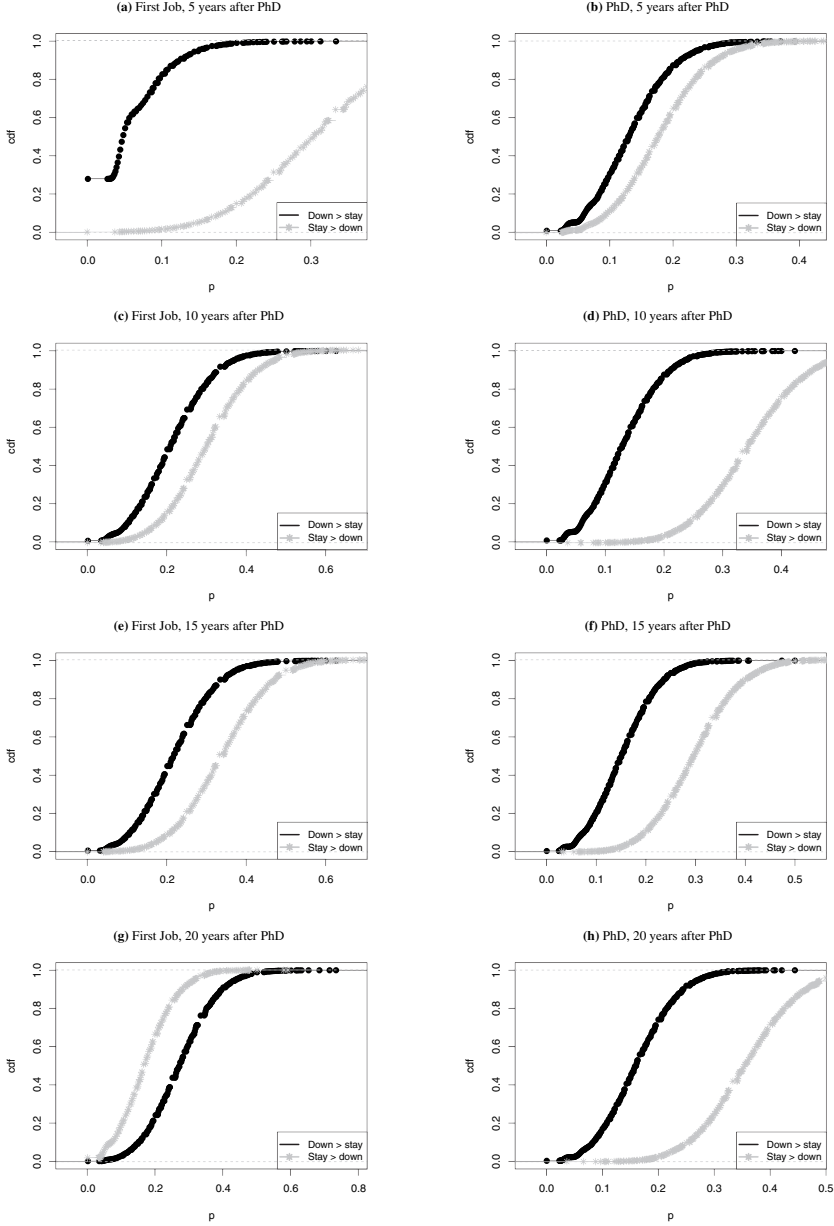
## **C.18 Cohort effects**

In this appendix I present the results of the matched pair analysis but restrict the sample to those who received a PhD prior to 1992. This way I reduce significantly any possible cohort effects.

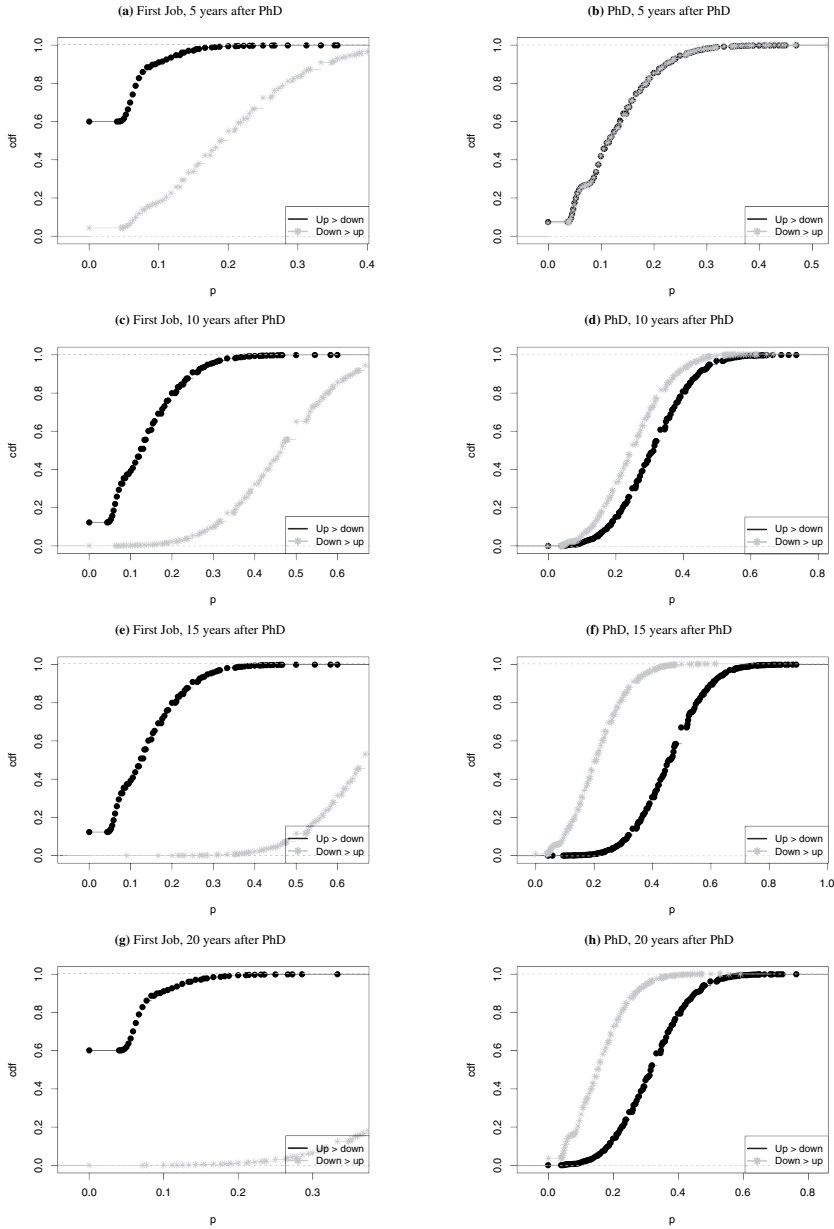




**Figure 15:** Up versus stay comparison, with the sample restricted to those who received their PhD degrees prior to 1992. The black curves are cumulative distribution functions of the proportion of observations in which  $R_{up} > R_{stay}$  was the case for  $p\%$  of the matched pairs. Grey curves are the CDFs for the  $R_{stay} > R_{up}$  proportions. From top to bottom 5, 10, 15, and 20 years after PhD. Pairs matched using gender, race, PhD obtained years and first job university (left column) or PhD institution (right column).



**Figure 16:** Down versus stay comparison, with the sample restricted to those who received their PhD degrees prior to 1992. The black curves are cumulative distribution functions of the proportion of observations in which  $R_{down} > R_{stay}$  was the case for  $p\%$  of the matched pairs. Grey curves are the CDFs for the  $R_{stay} > R_{down}$  proportions. From top to bottom 5,10,15, and 20 years after PhD. Pairs matched using gender, race, PhD obtained years and first job university (left column) or PhD institution (right column).



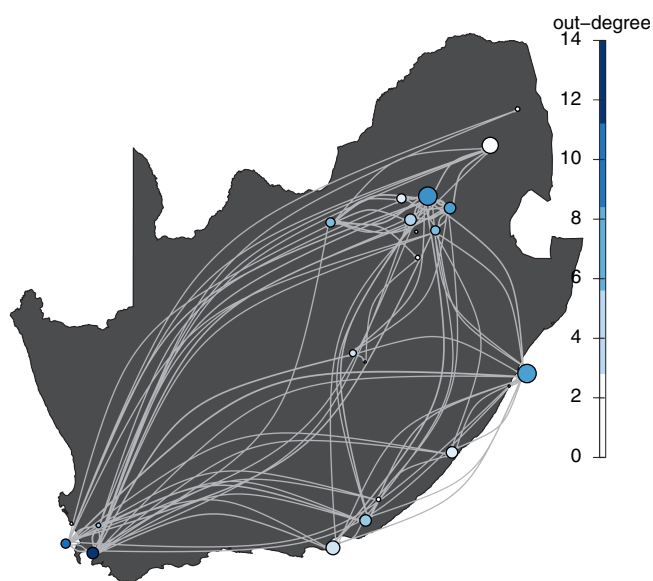
**Figure 17:** Up versus down comparison, with the sample restricted to those who received their PhD degrees prior to 1992. The black curves are cumulative distribution functions of the proportion of observations in which  $R_{up} > R_{down}$  was the case for  $p\%$  of the matched pairs. Grey curves are the CDFs for the  $R_{down} > R_{up}$  proportions. From top to bottom 5,10,15, and 20 years after PhD. Pairs matched using gender, race, PhD obtained years and first job university (left column) or PhD institution (right column).

## C.19 SSH results

Table 35 gives the adjacency matrix of PhD to first job transition in SSH.

PhD Institution		First Job Institution																									
		CapeTown	NelsonMandelaMetropolitan	Witwatersrand	Pretoria	Johannesburg	TheFreeState	SouthAfrica	Stellenbosch	KwaZuluNatal	TheNorthWest	Limpopo	Rhodes	TheWesternCape	FortHare	WalterSisulu	Venda	CentralUnivOfTechnology	TshwaneUnivOfTechnology	VaalUnivOfTechnology	MonashSA	DurbanInstituteOfTechnology	CapePeninsulaUnivOfTechnology				
	CapeTown	30	0	1	0	0	0	3	1	5	1	1	3	4	1	1	1	0	0	0	0	0	0	0	0	0	0
	NelsonMandelaMetropolitan	0	11	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	Witwatersrand	3	0	27	3	0	0	2	0	2	0	0	1	0	0	1	0	0	1	0	0	0	0	0	0	0	0
	Pretoria	0	1	3	48	5	1	10	0	0	4	4	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0
	Johannesburg	0	0	0	1	16	0	4	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	TheFreeState	0	0	0	0	0	24	0	0	1	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0
	SouthAfrica	1	4	0	11	1	2	69	0	3	5	3	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
	Stellenbosch	4	2	1	3	1	4	2	42	1	0	1	1	5	1	1	1	0	0	0	0	0	0	0	0	0	0
	KwaZuluNatal	1	2	1	0	0	0	1	1	36	0	0	1	0	0	2	0	0	0	0	0	0	0	2	0	0	0
	TheNorthWest	0	0	0	0	3	0	5	1	0	36	1	0	0	0	0	0	0	2	2	0	0	0	0	0	0	0
	Limpopo	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	Rhodes	0	1	0	0	1	0	3	0	2	0	0	7	0	0	1	0	0	0	0	0	0	0	0	0	0	0
	TheWesternCape	0	1	0	0	0	0	0	2	0	0	1	14	0	0	0	0	0	1	0	0	0	0	0	0	0	0
	FortHare	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	WalterSisulu	0	0	0	0	0	0	0	0	1	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0
	Venda	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	CentralUnivOfTechnology	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
	TshwaneUnivOfTechnology	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
	VaalUnivOfTechnology	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	MonashSA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	DurbanInstituteOfTechnology	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	CapePeninsulaUnivOfTechnology	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

**Table 35:** Adjacency matrix of the hiring network for the years 1970-2004 in SSH, rows are PhD institutions and columns are first job institutions. Each entry represents the number for people with a PhD in university  $i$  hired as first job in university  $j$ .



**Figure 18:** Hiring network 1970-2004 SSH. The vertex are the South African Universities, plotted according to their geographical coordinates (for the institutions located in the same area I separated manually). Vertex size in-degree, vertex colour out-degree. Where the correlation between in-degree and out-degree 0.53.

Table 36 contains summary statistics of the SSH hiring network.

### Summary Statistics SSH hiring network

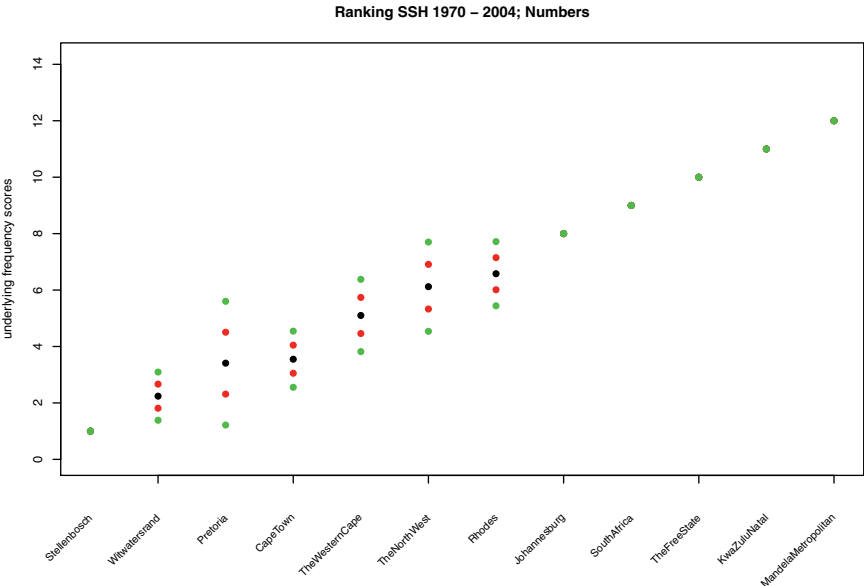
	All	Male	Female	White	Black
Number of Nodes	22	22	22	22	22
Number of Components	3	3	5	3	7
Number of isolated Nodes	2	2	4	2	6
Statistics on the Giant Component					
Number of Nodes	20	20	18	20	16
Number of Edges	83	63	43	74	26
Edge Density	0.218	0.166	0.141	0.195	0.108
Average Path Length	1.959	1.976	2.386	2.02	2.037
Diameter	8	11	7	9	5
Global Clustering Coefficient	0.525	0.429	0.35	0.51	0.308

**Table 36:** Summary Statistics SSH hiring network for the years 1970-2004. Network statistics are compute without considering self-loops.

### SSH PhDs hired from the top 5 universities

Rank	PhD University	Placed in SA academia	Hired by Top 5	Proportion placed in Top 5
1	Stellenbosch	70	55	0.786
2	Witwatersrand	39	33	0.846
3	Cape Town	52	36	0.692
4	Pretoria	79	51	0.646
5	WesternCape	20	16	0.8
	All in Top 5	260	191	0.735
Total PhD hires		542		0.480

**Table 37:** PhDs hired from the top 5 prestigious universities in SSH, according to the prestige ranking. The total of SSH PhDs Hired in the period is 542.



**Figure 19:** Prestige Ranking for SSH 1970-2004. The frequency scores are in ascending order from the highest prestige which correspond to one. The black dots is the average of the orders with the maximum scores under 10000 repetition, red dots and green dots are respectively one and two standard deviation from the average. The algorithm runs on the adjacency matrix of the hiring network. Universities with fewer than 5 PhDs are excluded.

## C.20 KS test matched pairs

## Results KS test matched pairs analysis

Up-stay			Down-stay			Up-Down		
	First Job		First Job		PHD	First Job		PHD
	PHD		PHD			PHD		
two tailed less greater	D=0.97; p-value<2.10 <sup>-16</sup>		5 years after PHD					
	D=0.61; p-value<2.10 <sup>-16</sup>		D=0.32; p-value<2.10 <sup>-16</sup>	D=0.15; p-value<2.10 <sup>-16</sup>		D=0.10; p-value<2.10 <sup>-16</sup>	D=0.77; p-value<2.10 <sup>-16</sup>	
	D=0.97; p-value<2.10 <sup>-16</sup>	D=0; p-value=1	D=0; p-value=1	D=0.15; p-value<2.10 <sup>-16</sup>		D=0; p-value=1	D=0.77; p-value<2.10 <sup>-16</sup>	
	D=0; p-value=1		D=0.32; p-value<2.10 <sup>-16</sup>	D=0; p-value=1		D=0.10; p-value<2.10 <sup>-16</sup>	D=0; p-value=1	
two tailed less greater	D=0.02; p-value=0.013		10 years after PHD					
	D=7.10 <sup>-4</sup> ; p-value=0.995		D=0.30; p-value<2.10 <sup>-16</sup>	D=0.96; p-value<2.10 <sup>-16</sup>		D=0.23; p-value<2.10 <sup>-16</sup>	D=0.54; p-value<2.10 <sup>-16</sup>	
	D=0.02; p-value=0.007		D=0; p-value=1	D=0; p-value=1		D=0; p-value=1	D=0.54; p-value<2.10 <sup>-16</sup>	
			D=0; p-value=1	D=0.96; p-value<2.10 <sup>-16</sup>		D=0.23; p-value<2.10 <sup>-16</sup>	D=0; p-value=1	
two tailed less greater	D=0.21; p-value<2.10 <sup>-16</sup>		15 years after PHD					
	D=0; p-value=1		D=0.64; p-value<2.10 <sup>-16</sup>	D=0.80; p-value<2.10 <sup>-16</sup>		D=0.95; p-value<2.10 <sup>-16</sup>	D=0.81; p-value<2.10 <sup>-16</sup>	
	D=0.21; p-value<2.10 <sup>-16</sup>	D=0; p-value=1	D=0; p-value=1	D=0; p-value=1		D=0; p-value=1	D=0.81; p-value<2.10 <sup>-16</sup>	
			D=0.64; p-value<2.10 <sup>-16</sup>	D=0.80; p-value<2.10 <sup>-16</sup>		D=0.95; p-value<2.10 <sup>-16</sup>	D=0; p-value=1	
two tailed less greater	D=0.51; p-value<2.10 <sup>-16</sup>		20 years after PHD					
	D=0; p-value=1		D=0.39; p-value<2.10 <sup>-16</sup>	D=0.76; p-value<2.10 <sup>-16</sup>		D=0.97; p-value<2.10 <sup>-16</sup>	D=0.58; p-value<2.10 <sup>-16</sup>	
	D=0.51; p-value<2.10 <sup>-16</sup>	D=0.79; p-value<2.10 <sup>-16</sup>	D=0.39; p-value<2.10 <sup>-16</sup>	D=0; p-value=1		D=0; p-value=1	D=0.58; p-value<2.10 <sup>-16</sup>	
			D=0.79; p-value<2.10 <sup>-16</sup>	D=0; p-value=1		D=0.97; p-value<2.10 <sup>-16</sup>	D=0; p-value=1	

**Table 38:** Results of KS test relative to figure 4.3, 4.4, and 4.5. For two tailed test the null hypothesis is that cdf are different, for the less/greater test the first cdf lies below/above the other.



## Results KS test matched pairs analysis; Cohort effects

	Up-stay			Down-stay			Up-Down		
	First Job			First Job			First Job		
	PhD			PhD			PhD		
two tailed less greater	D=0.25; p-value<2.2.10 <sup>-16</sup>			5 years after PhD					
	D=0.25; p-value<2.2.10 <sup>-16</sup>	D=0.66; p-value<2.2.10 <sup>-16</sup>	D=0.90; p-value<2.2.10 <sup>-16</sup>	D=0.30; p-value<2.2.10 <sup>-16</sup>	D=0.73; p-value<2.2.10 <sup>-16</sup>	D=0.007; p-value=0.97			
	D=0; p-value=1	D=0.66; p-value<2.2.10 <sup>-16</sup>	D=0; p-value=1	D=0; p-value=1	D=0; p-value=1	D=0.007; p-value=0.64			
two tailed less greater	D=0.27; p-value<2.2.10 <sup>-16</sup>	D=0.06; p-value<2.2.10 <sup>-16</sup>	D=0.37; p-value<2.2.10 <sup>-16</sup>	10 years after PhD					
	D=0.27; p-value<2.2.10 <sup>-16</sup>	D=0; p-value=1	D=0; p-value=1	D=0.86; p-value<2.2.10 <sup>-16</sup>	D=0.86; p-value<2.2.10 <sup>-16</sup>	D=0.24; p-value<2.2.10 <sup>-16</sup>			
	D=0.27; p-value<2.2.10 <sup>-16</sup>	D=0.06; p-value<2.2.10 <sup>-16</sup>	D=0.37; p-value<2.2.10 <sup>-16</sup>	D=0; p-value=1	D=0.86; p-value<2.2.10 <sup>-16</sup>	D=0.24; p-value<2.2.10 <sup>-16</sup>			
two tailed less greater	D=0.45; p-value<2.2.10 <sup>-16</sup>	D=0.69; p-value<2.2.10 <sup>-16</sup>	D=0.47; p-value<2.2.10 <sup>-16</sup>	15 years after PhD					
	D=0; p-value=1	D=0; p-value=1	D=0; p-value=1	D=0.69; p-value<2.2.10 <sup>-16</sup>	D=0.98; p-value<2.2.10 <sup>-16</sup>	D=0.77; p-value<2.2.10 <sup>-16</sup>			
	D=0.45; p-value<2.2.10 <sup>-16</sup>	D=0.69; p-value<2.2.10 <sup>-16</sup>	D=0.47; p-value<2.2.10 <sup>-16</sup>	D=0; p-value=1	D=0; p-value=1	D=0.24; p-value<2.2.10 <sup>-16</sup>			
two tailed less greater	D=0.53; p-value<2.2.10 <sup>-16</sup>	D=0.82; p-value<2.2.10 <sup>-16</sup>	D=0.43; p-value<2.2.10 <sup>-16</sup>	20 years after PhD					
	D=0.53; p-value<2.2.10 <sup>-16</sup>	D=0; p-value=1	D=0.43; p-value<2.2.10 <sup>-16</sup>	D=0.82; p-value<2.2.10 <sup>-16</sup>	D=0.98; p-value<2.2.10 <sup>-16</sup>	D=0.61; p-value<2.2.10 <sup>-16</sup>			
	D=0.53; p-value<2.2.10 <sup>-16</sup>	D=0.82; p-value<2.2.10 <sup>-16</sup>	D=0.43; p-value<2.2.10 <sup>-16</sup>	D=0; p-value=1	D=0; p-value=1	D=0.61; p-value<2.2.10 <sup>-16</sup>			

**Table 39:** Results of KS test relative to figures 15, 16, and 17. Sample restricted to those who received their PhD degrees prior to 1992 . For two tailed test the null hypothesis is that cdf are different, for the less/greater test the first cdf lies below/above the other.

## C.21 Robustness check of KS test

To address the well known problem of excess of sensitivity of KS test with large samples, we perform a robustness check, following a bootstrap technique. For each comparison of the distributions of proportions obtained with our matched pairs technique, we sample with replacement 100 samples of size 100 and each time we compute the KS test<sup>3</sup>, then we store the obtained p-values and we count how many times the p-values are lower than 0.05. Under the null hypothesis, p-values are distributed as a uniform, so if the fraction of p-values under 0.05 is larger than 0.05, then we can conclude in a more consistent way that the two distributions are different. Table 40 shows the results of this procedure for figures 4.3, 4.4, and 4.5. Observing the table the results of KS tests are confirmed.

---

<sup>3</sup>According to Sekhon (2011) to account for possible presence of ties we use the `ks.boot` command under the R library `Matching`.

### Matched pairs robustness analysis KS test

	Up-stay		Down-stay		Up-Down	
	First Job	PhD	First Job	PhD	First Job	PhD
	5 years after PhD					
two tailed	1	1	1	0.54	0.28	1
less	1	1	0	0.65	0	1
greater	0	0	1	0	0.42	0
	10 years after PhD					
two tailed	0.03	0.26	0.99	1	0.94	1
less	0	0.38	0	0	0	1
greater	0.06	0	0.99	1	0.96	0
	15 years after PhD					
two tailed	0.89	1	1	1	1	1
less	0	0	0	0	0	1
greater	0.94	1	1	1	1	0
	20 years after PhD					
two tailed	1	1	1	1	1	1
less	0	0	1	0	0	1
greater	1	1	0	1	1	0

**Table 40:** Proportion of p-values lower than 0.05 of Bootstrap KS test repeated 100 times on repeated re-samples of size=100 for p%

### Matched pairs robustness analysis KS test; Cohort effects

	Up-stay		Down-stay		Up-Down	
	First Job	PhD	First Job	PhD	First Job	PhD
5 years after PhD						
two tailed	0.94	1	1	0.99	1	0.05
less	0.96	1	0	0	0	0.03
greater	0	0	1	0.99	1	0.07
10 years after PhD						
two tailed	0.96	0.14	1	1	1	0.98
less	0	0	0	0	0	0.98
greater	0.99	0.21	1	1	1	0
15 years after PhD						
two tailed	1	1	1	1	1	1
less	0	0	0	0	0	1
greater	1	1	1	1	1	0
20 years after PhD						
two tailed	1	1	1	1	1	1
less	0	0	1	0	0	1
greater	1	1	0	1	1	0

**Table 41:** Proportion of p-values lower than 0.05 of Bootstrap KS test repeated 100 times on repeated re-samples of size=100 for p%. Sample restricted to those who received their PhD degrees prior than 1992



## D.22 Faculty Hiring Matrix Names

	ABBREVIATION	NAMES
1	CINVESTAV(ZACATENCO)	CENTRO DE INVESTIGACION Y DE ESTUDIOS AVANZADOS DEL I.P.N. - UNIDAD ZACATENCO
2	CICESE	CENTRO DE INVESTIGACION CIENTIFICA Y DE EDUCACION SUPERIOR DE ENSENADA
3	UNAM	UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO
4	CINVESTAV(OTHERS)	CENTRO DE INVESTIGACION Y DE ESTUDIOS AVANZADOS DEL INSTITUTO POLITECNICO NACIONAL
5	IMP	INSTITUTO MEXICANO DEL PETROLEO
6	ECOSUR	EL COLEGIO DE LA FRONTERA SUR
7	IPICYT	INSTITUTO POTOSINO DE INVESTIGACION CIENTIFICA Y TECNOLOGICA AC
8	CIBINOR	CENTRO DE INVESTIGACIONES BIOLOGICAS DEL NOROESTE SC
9	INAOE	INSTITUTO NACIONAL DE ASTROFISICA OPTICA Y ELECTRONICA
10	UAM	UNIVERSIDAD AUTONOMA METROPOLITANA
11	UAEM	UNIVERSIDAD AUTONOMA DEL ESTADO DE MORELOS
12	INECOL	INSTITUTO DE ECOLOGIA, A.C
13	IBERO	UNIVERSIDAD IBEROAMERICANA AC
14	CICY	CENTRO DE INVESTIGACION CIENTIFICA DE YUCATAN AC
15	CIO	CENTRO DE INVESTIGACIONES EN OPTICA AC
16	COLPOS	COLEGIO DE POSTGRADUADOS
17	CIQA	CENTRO DE INVESTIGACION EN QUIMICA APLICADA, A.C
18	CIAD	CENTRO DE INVESTIGACION EN ALIMENTACION Y DESARROLLO AC
19	BUAP	BENEMERITA UNIVERSIDAD AUTONOMA DE PUEBLA
20	UASLP	UNIVERSIDAD AUTONOMA DE SAN LUIS POTOSI
21	CIMAV	CENTRO DE INVESTIGACION EN MATERIALES AVANZADOS SC
22	IPN	INSTITUTO POLITECNICO NACIONAL
23	UV	UNIVERSIDAD VERACRUZANA
24	UADY	UNIVERSIDAD AUTONOMA DE YUCATAN
25	UAEM	UNIVERSIDAD AUTONOMA DEL ESTADO DE MEXICO
26	UCOL	UNIVERSIDAD DE COLIMA
27	UAQ	UNIVERSIDAD AUTONOMA DE QUERETARO
28	TEC	INSTITUTO TECNOLÓGICO Y DE ESTUDIOS SUPERIORES DE MONTERREY
29	UDG	UNIVERSIDAD DE GUADALAJARA
30	UGTO	UNIVERSIDAD DE GUANAJUATO
31	IMSS	INSTITUTO MEXICANO DEL SEGURO SOCIAL
32	UNISON	UNIVERSIDAD DE SONORA
33	UMICH	UNIVERSIDAD MICHOACANA DE SAN NICOLAS DE HIDALGO
34	UAEH	UNIVERSIDAD AUTONOMA DEL ESTADO DE HIDALGO
35	UANL	UNIVERSIDAD AUTONOMA DE NUEVO LEON
36	TECNM	TECNOLÓGICO NACIONAL DE MEXICO

## D.23 Prestige Ranking Algorithm

```

/* Input:  A:= weighted directed Adjacency Matrix with zero
           entries in the main diagonal.                                */
sk:= A top-score defined with an arbitrary small initial value;
S:= An empty vector of max scores;
O:= An empty matrix of orders with u columns;
for  $i = 0$  to  $n$  do
    o0:= Generate a random n-tuple of length u;
    M:= Sort the matrix A by o0;
    s0:= Compute current score by adding the upper triangular elements of M
         by row;
    /* Local Search                                                    */
    for  $i = 0$  to  $n$  do
        o1:= Swap randomly two elements of o0;
        M:= Sort the matrix M by o1;
        s1:= Compute swap score by adding the upper triangular elements of
             M by row;
        if  $s1 \geq s0$  then
            M:= Sort the matrix A by o1;
            sk:= Update the top-score by s1;
        end if
    end for
    dif:= Compute difference between s1 and sk;
    if  $dif > 0$  then
        sk:= Update the top-score by s1;
        S:= Add the current score sk as an element of S ;
        O:= Add a row of orders given by s1 ;
    end if
end for
/* Output1:  S:= A set of highest scores                                */
/* Output2:  O:= A matrix of orders for each element in S
*/

```

**Algorithm 1:** Prestige Ranking Algorithm by Cowan and Rossello (2018).

## D.24 NSR: Disciplines, Evaluation and Commissions

### Disciplines

Area I	Physics-Mathematics and Earth Sciences
Area II	Biology, Chemistry, Life Sciences
Area III	Medicine and Health Sciences
Area IV	Humanities and Behavioural Sciences
Area V	Social Sciences
Area VI	Biotechnology and Agricultural Sciences
Area VII	Engineering

### Commissions

The members of the "Evaluation Commission" are designated by the "Council of Approval", the highest authority in the "National System of Researchers" (NSR). Once the Evaluation Commission is designated and formalised it will run for a period of 3 years. The Evaluation Commission has the obligation to review applications of new and incumbent members of the NSR.

### Evaluation

The "Evaluation Commission" reviews each application assessing both the quality and the quantity of the research output. Each application is evaluated by at least two members of the commission. The evaluation takes into consideration primarily the research output, but also human capital formation (number of supervisions) and linkages with industry and the public sector, updates in study plans and publication of dissemination articles. The research output include: Articles, Books, Book Chapters, Patents, Technological Developments, Innovations and Transfers of Technology.

### Levels of Rewards:

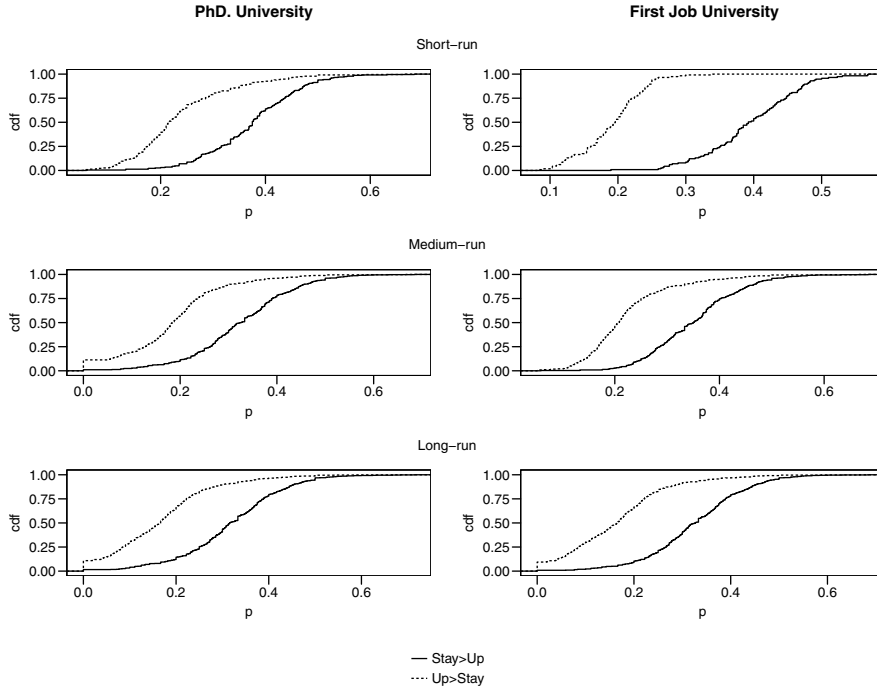
- SNI Candidates<sup>4</sup>: Granted for 3 years, with the possibility of 2 years of extension.
- Level I: Granted for 3 years the first time, and every 4 years in the following periods.
- Level II: Granted for 4 years the first time, and every 5 years in the following periods.
- Level III: Granted for 5 years the first and second time, and every 10 years in the following periods.
- Emeritus Professors: Candidates must have 65 years of more, and have accumulated at least three periods of level III distinction (15 years) without interruption.

---

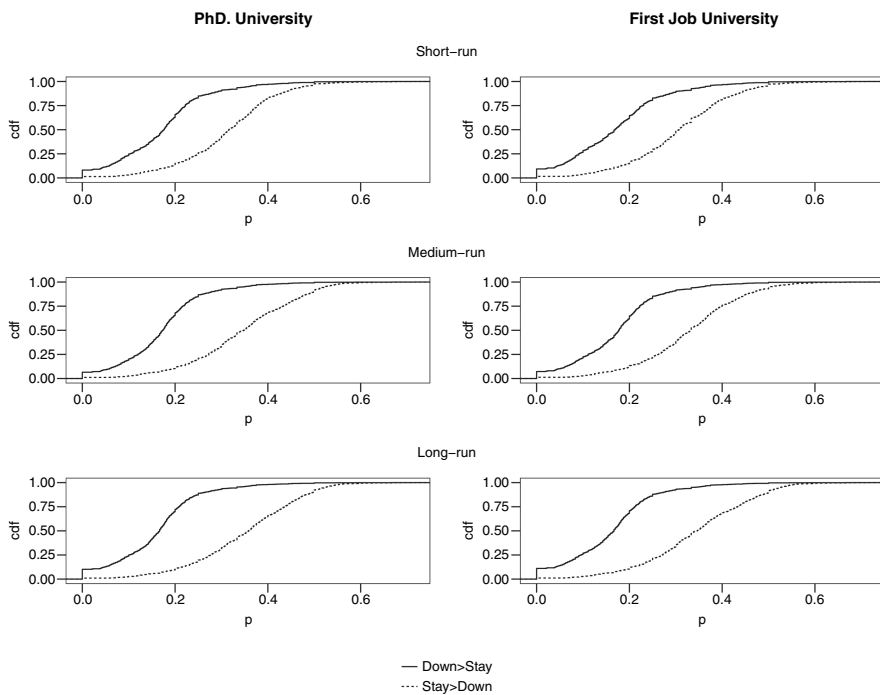
<sup>4</sup>Applicants can only receive this distinction one time.



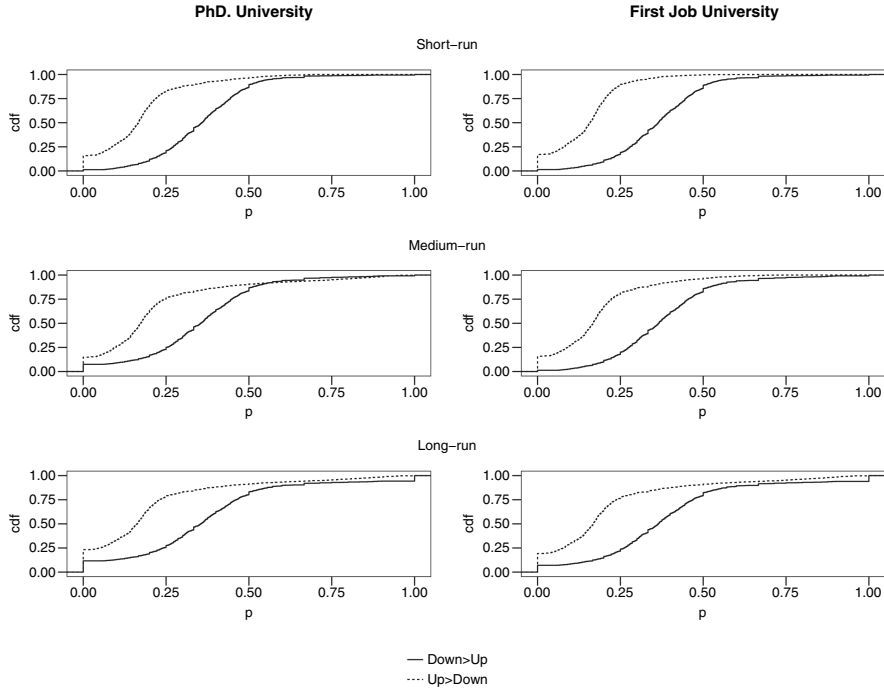
## D.25 Results Static Rank



**Figure 20:** Stochastic Analysis of Up vs Stay. The solid curves are CEDFs of the proportion of pairs in which  $R_{Stay} > R_{Up}$ . Dotted curves are CEDFs for  $R_{Up} > R_{Stay}$ . Pairs matched by gender, age, discipline, graduation years, and same PhD university (left), or same first-job university (right). From top to bottom: short-run (Up to 2 years), medium-run (3-5 years), and long-run (6-25 years) after PhD graduation.



**Figure 21: Stochastic Analysis of Stay vs Down.** The solid curves are CEDFs of the proportion of pairs in which  $R_{Down} > R_{Stay}$ . Dotted curves are CEDFs for  $R_{Stay} > R_{Down}$ . Pairs matched by gender, age, discipline, graduation years, and same PhD university (left), or same first-job university (right). From top to bottom: short-run (Up to 2 years), medium-run (3-5 years), and long-run (6-25 years) after PhD graduation.



**Figure 22:** Stochastic Analysis of **Stay vs Down**. The solid curves are CEDFs of the proportion of pairs in which  $R_{Down} > R_{Stay}$ . Dotted curves are CEDFs for  $R_{Stay} > R_{Down}$ . Pairs matched by gender, age, discipline, graduation years, and same PhD university (left), or same first-job university (right). From top to bottom: short-run (Up to 2 years), medium-run (3-5 years), and long-run (6-25 years) after PhD graduation.

## D.26 KS-Test Results

**Table 42:** KS-Test Results for Phd and First-job Matches.

PhD Matches			
$Null : H_0$	Short-run	Med-run	Long-run
$F(p^{Up}) = F(p^{Stay})$	D=0.9874; pvalue=0	D=0.9097; pvalue=0	D=0.8165; pvalue=0
$F(p^{Up}) \leq F(p^{Stay})$	D=0; pvalue=1	D=0; pvalue=1	D=0; pvalue=1
$F(p^{Up}) \geq F(p^{Stay})$	D=0.9874; pvalue=0	D=0.9097; pvalue=0	D=0.8165; pvalue=0
$F(p^{Down}) = F(p^{Stay})$	D=0.812; pvalue=0	D=0.8487; pvalue=0	D=0.8564; pvalue=0
$F(p^{Down}) \leq F(p^{Stay})$	D=0; pvalue=1	D=0; pvalue=1	D=0; pvalue=1
$F(p^{Down}) \geq F(p^{Stay})$	D=0.812; pvalue=0	D=0.8487; pvalue=0	D=0.8564; pvalue=0
$F(p^{Up}) = F(p^{Down})$	D=0.7353; pvalue=0	D=0.6077; pvalue=0	D=0.5924; pvalue=0
$F(p^{Up}) \leq F(p^{Down})$	D=0; pvalue=1	D=0; pvalue=1	D=0; pvalue=1
$F(p^{Up}) \geq F(p^{Down})$	D=0.7353; pvalue=0	D=0.6077; pvalue=0	D=0.5924; pvalue=0

First-Job Matches			
$Null : H_0$	Short-run	Med-run	Long-run
$F(p^{Up}) = F(p^{Stay})$	D=0.9227; pvalue=0	D=0.888; pvalue=0	D=0.9227; pvalue=0
$F(p^{Up}) \leq F(p^{Stay})$	D=0; pvalue=1	D=0; pvalue=1	D=0; pvalue=1
$F(p^{Up}) \geq F(p^{Stay})$	D=0.9227; pvalue=0	D=0.888; pvalue=0	D=0.9227; pvalue=0
$F(p^{Down}) = F(p^{Stay})$	D=0.8434; pvalue=0	D=0.8319; pvalue=0	D=0.8434; pvalue=0
$F(p^{Down}) \leq F(p^{Stay})$	D=0; pvalue=1	D=0; pvalue=1	D=0; pvalue=1
$F(p^{Down}) \geq F(p^{Stay})$	D=0.8434; pvalue=0	D=0.8319; pvalue=0	D=0.8434; pvalue=0
$F(p^{Up}) = F(p^{Down})$	D=0.6273; pvalue=0	D=0.6835; pvalue=0	D=0.6273; pvalue=0
$F(p^{Up}) \leq F(p^{Down})$	D=0; pvalue=1	D=0; pvalue=1	D=0; pvalue=1
$F(p^{Up}) \geq F(p^{Down})$	D=0.6273; pvalue=0	D=0.6835; pvalue=0	D=0.6273; pvalue=0

---

# Bibliography

- Abbott, P. and Malunda, D. (2016). The promise and the reality: Women's rights in Rwanda. *African Journal of International and Comparative Law*, 24(4):561–581.
- Abbott, W. F. and Barlow, H. M. (1972). Stratification theory and organizational rank: Resources, functions, and organizational prestige in american universities. *Pacific Sociological Review*, 15(4):401–424.
- Aguinis, H., Ji, Y. H., and Joo, H. (2018). Gender productivity gap among star performers in STEM and other scientific fields. *Journal of Applied Psychology*, 103(12):1283.
- Akerlind, G. (2009). Postdoctoral research positions as preparation for an academic career. *International Journal for Researcher Development*, 1(1):84–96.
- Alderfer, C. P. and Smith, K. K. (1982). Studying intergroup relations embedded in organizations. *Administrative Science Quarterly*, pages 35–65.
- Allen, S. G. (2018). Black and Hispanic students get lower return on higher education investment. *Diverse Issues in Higher Education*, 35(10):10–10.
- Allison, P. D. and Long, J. S. (1987). Inter-university mobility of academic scientists. *American Sociological Review*, pages 643–652.
- Allison, P. D. and Long, J. S. (1990). Departmental effects on scientific productivity. *American Sociological Review*, 55(4):469–478.
- Allison, P. D. and Stewart, J. A. (1974). Productivity differences among scientists: Evidence for accumulative advantage. *American Sociological Review*, 39(4):596–606.
- Allport, G. W., Clark, K., and Pettigrew, T. (1954). The nature of prejudice.
- Araki, S., Kawaguchi, D., and Onozuka, Y. (2016). University prestige, performance evaluation, and promotion: Estimating the employer learning model using personnel datasets. *Labour Economics*, 41:135–148.
- Astin, A. W. (1969). Comment on. A student's dilemma: Big fish-little pond or little fish-big pond.

- 
- Astin, A. W. and Astin, H. S. (2000). Leadership reconsidered: Engaging higher education in social change.
- Azoulay, P., Graff Zivin, J. S., and Manso, G. (2011). Incentives and creativity: evidence from the academic life sciences. *The RAND Journal of Economics*, 42(3):527–554.
- Azoulay, P., Stuart, T., and Wang, Y. (2014). Matthew: Effect or fable? *Management Science*, 60(1):92–109.
- Azoulay, P., Zivin, J. S. G., and Sampat, B. N. (2012). *The diffusion of scientific knowledge across time and space: Evidence from professional transitions for the superstars of medicine*, pages 107–155. University of Chicago Press, 107–155, Chicago IL.
- Bagues, M., Sylos-Labini, M., and Zinovyeva, N. (2014). Do gender quotas pass the test? evidence from academic evaluations in Italy. *Scuola Superiore Sant’Anna, LEM Working Paper Series*, 14.
- Bagues, M., Sylos-Labini, M., and Zinovyeva, N. (2017). Does the gender composition of scientific committees matter? *American Economic Review*, 107(4):1207–38.
- Bagues, M. F. and Esteve-Volart, B. (2010). Can gender parity break the glass ceiling? evidence from a repeated randomized experiment. *The Review of Economic Studies*, 77(4):1301–1328.
- Bair, J. H. (2003). Hiring practices in finance education: Linkages among top-ranked graduate programs. *American Journal of Economics and Sociology*, 62(2):429–433.
- Bair, J. H. and Bair, R. K. (1998). Linkages among top-ranked graduate programs in special education: A brief report. *Mental Retardation*, 36(1):52–54.
- Bair, J. H. and Boor, M. (1991). The academic elite in law: Linkages among top-ranked law schools. *Psychological Reports*, 68(3):891–894.
- Baldi, S. (1995). Prestige determinants of first academic job for new sociology Ph.Ds 1985–1992. *The Sociological Quarterly*, 36(4):777–789.
- Barabási, A.-L. and Albert, R. (1999). Emergence of scaling in random networks. *Science*, 286(5439):509–512.
- Barnard, H., Cowan, R., and Müller, M. (2012). Global excellence at the expense of local diffusion, or a bridge between two worlds? research in science and technology in the developing world. *Research Policy*, 41(4):756–769.
- Barnard, H., Cowan, R. A., Kirman, A. P., and Müller, M. (2016). Including excluded groups: The slow racial transformation of the South African university system. Technical report, Working Paper Series in Economics, Karlsruher Institut für Technologie (KIT).

- 
- Barnett, G. A. and Benefield, G. A. (2017). Predicting international Facebook ties through cultural homophily and other factors. *New Media & Society*, 19(2):217–239.
- Barnett, G. A., Danowski, J. A., Feeley, T. H., and Stalker, J. (2010). Measuring quality in communication doctoral education using network analysis of faculty hiring patterns. *Journal of Communication*, 60(2):388–411.
- Bazeley, P. (2003). Defining “early career” in research. *Higher Education*, 45(3):257–279.
- Becker, G. S. (1957). The economics of discrimination.
- Bedeian, A. G., Cavazos, D. E., Hunt, J. G., and Jauch, L. R. (2010). Doctoral degree prestige and the academic marketplace: A study of career mobility within the management discipline. *Academy of Management Learning & Education*, 9(1):11–25.
- Bedeian, A. G. and Feild, H. S. (1980). Academic stratification in graduate management programs: Departmental prestige and faculty hiring patterns. *Journal of Management*, 6(2):99–115.
- Bellou, A. and Cardia, E. (2016). Occupations after WWII: The legacy of rosie the riveter. *Explorations in Economic History*, 62:124–142.
- Bergeron, L. and Gordon, M. (2017). Establishing a STEM pipeline: Trends in male and female enrollment and performance in higher level secondary STEM courses. *International Journal of Science and Mathematics Education*, 15(3):433–450.
- Bettinger, E. P. and Long, B. T. (2005). Do faculty serve as role models? the impact of instructor gender on female students. *American Economic Review*, 95(2):152–157.
- Bian, L., Leslie, S.-J., and Cimpian, A. (2017). Gender stereotypes about intellectual ability emerge early and influence children’s interests. *Science*, 355(6323):389–391.
- Bickel, P. J., Hammel, E. A., and O’Connell, J. W. (1975). Sex bias in graduate admissions: Data from Berkeley. *Science*, 187(4175):398–404.
- Bidwell, M. (2011). Paying more to get less: The effects of external hiring versus internal mobility. *Administrative Science Quarterly*, 56(3):369–407.
- Blackburn, R. T., Chapman, D. W., and Cameron, S. M. (1981). Cloning in academe: Mentorship and academic careers. *Research in Higher Education*, 15(4):315–327.
- Blair, P. Q. and Deming, D. J. (2020). Structural increases in demand for skill after the great recession. In *AEA Papers and Proceedings*, volume 110, pages 362–65.
- Blalock, S. D. and Sharpe, R. V. (2012). You go girl!: Trends in educational attainment of black women. In *Black female undergraduates on campus: Successes and challenges*. Emerald Group Publishing Limited.



- 
- Blau, F. D. and Kahn, L. M. (2017). The gender wage gap: Extent, trends, and explanations. *Journal of Economic Literature*, 55(3):789–865.
- Boehnke, J. and Gay, V. (2020). The missing men World War I and female labor force participation. *Journal of Human Resources*, pages 0419–10151R1.
- Bol, T., de Vaan, M., and van de Rijt, A. (2018). The Matthew effect in Science funding. *Proceedings of the National Academy of Sciences*, 115(19):4887–4890.
- Boudreau, K. J., Guinan, E. C., Lakhani, K. R., and Riedl, C. (2016). Looking across and looking beyond the knowledge frontier: Intellectual distance, novelty, and resource allocation in science. *Management Science*, 62(10):2765–2783.
- Bramoullé, Y., Currarini, S., Jackson, M. O., Pin, P., and Rogers, B. W. (2012). Homophily and long-run integration in social networks. *Journal of Economic Theory*, 147(5):1754–1786.
- Breda, T., Grenet, J., Monnet, M., and Van Effenterre, C. (2018). Can female role models reduce the gender gap in science? evidence from classroom interventions in french high schools. *PSE Working Papers halshs-01713068*.
- Breetzke, G. D. and Hedding, D. W. (2018). The changing demography of academic staff at higher education institutions (HEIs) in South Africa. *Higher Education*, 76(1):145–161.
- Brief, A. P., Umphress, E. E., Dietz, J., Burrows, J. W., Butz, R. M., and Scholten, L. (2005). Community matters: Realistic group conflict theory and the impact of diversity. *Academy of Management Journal*, 48(5):830–844.
- Brink, M. and Benschop, Y. (2014). Gender in academic networking: The role of gatekeepers in professorial recruitment. *Journal of Management Studies*, 51(3):460–492.
- Buela-Casal, G., Bermúdez, M. P., Sierra, J. C., Quevedo-Blasco, R., Castro, Á., and Guillén-Riquelme, A. (2012). Ranking de 2011 en producción y productividad en investigación de las universidades públicas españolas. *Psicothema*, 24(4):505–515.
- Burris, V. (2004). The academic caste system: Prestige hierarchies in Ph.D exchange networks. *American Sociological Review*, 69(2):239–264.
- Byrd, R. H., Lu, P., Nocedal, J., and Zhu, C. (1995). A limited memory algorithm for bound constrained optimization. *SIAM Journal on Scientific Computing*, 16(5):1190–1208.
- Chan, K. C., Chen, C. R., and Steiner, T. L. (2002). Production in the finance literature, institutional reputation, and labor mobility in academia: A global perspective. *Financial Management*, pages 131–156.

- 
- Chen, C., Chen, Y., Horowitz, M., Hou, H., Liu, Z., and Pellegrino, D. (2009). Towards an explanatory and computational theory of scientific discovery. *Journal of Informetrics*, 3(3):191–209.
- Cheryan, S., Ziegler, S. A., Montoya, A. K., and Jiang, L. (2017). Why are some STEM fields more gender balanced than others? *Psychological Bulletin*, 143(1):1.
- Chuang, K.-Y. and Ho, Y.-S. (2014). Bibliometric profile of top-cited single-author articles in the science citation index expanded. *Journal of Informetrics*, 8(4):951–962.
- Clauset, A., Arbesman, S., and Larremore, D. B. (2015). Systematic inequality and hierarchy in faculty hiring networks. *Science Advances*, 1(1):e1400005.
- Cole, J. R. and Zuckerman, H. (1984). The productivity puzzle. *Advances in Motivation and Achievement. Women in Science*.
- Confraria, H. and Godinho, M. M. (2015). The impact of African science: A bibliometric analysis. *Scientometrics*, 102(2):1241–1268.
- Connelly, B. L., Certo, S. T., Ireland, R. D., and Reutzel, C. R. (2011). Signaling theory: A review and assessment. *Journal of Management*, 37(1):39–67.
- Conti, A. and Visentin, F. (2015a). A revealed preference analysis of PhD students choices over employment outcomes. *Research Policy*, 44(10):1931–1947.
- Conti, A. and Visentin, F. (2015b). Science and engineering Ph.D. students’ career outcomes, by gender. *PloS One*, 10(8):e0133177.
- Cowan, R. and Rossello, G. (2018). Emergent structures in faculty hiring networks, and the effects of mobility on academic performance. *Scientometrics*, 117(1):527–562.
- Cowan, W. B., Cowan, R., and Llerena, P. (2010). Running the marathon. In McKelvey, M. and Holmén, M., editors, *Learning to Compete in European Universities From Social Institution to Knowledge Business*, chapter 11, pages 278–299. Edward Elgar Publishing.
- Crane, D. (1970). The academic marketplace revisited: A study of faculty mobility using the carter ratings. *American Journal of Sociology*, 75(6):953–964.
- Crisp, G. and Cruz, I. (2009). Mentoring college students: A critical review of the literature between 1990 and 2007. *Research in Higher Education*, 50(6):525–545.
- Crosby, J. R. and Monin, B. (2007). Failure to warn: How student race affects warnings of potential academic difficulty. *Journal of Experimental Social Psychology*, 43(4):663–670.

- 
- Currarini, S., Jackson, M. O., and Pin, P. (2009). An economic model of friendship: Homophily, minorities, and segregation. *Econometrica*, 77(4):1003–1045.
- Cyrenne, P. and Grant, H. (2009). University decision making and prestige: An empirical study. *Economics of Education Review*, 28(2):237–248.
- David, P. A. (1993). *Positive feedbacks and research productivity in science: Reopening another black box*. MERIT Buxton.
- De Paola, M., Ponzio, M., and Scoppa, V. (2015). Gender differences in attitudes towards competition: Evidence from the Italian scientific qualification.
- Debackere, K. and Rappa, M. A. (1995). Scientists at major and minor universities: Mobility along the prestige continuum. *Research Policy*, 24(1):137–150.
- Demeter, M. and Toth, T. (2020). The world-systemic network of global elite sociology: the western male monoculture at faculties of the top one-hundred sociology departments of the world. *Scientometrics*, pages 1–27.
- Edström, A. and Galbraith, J. R. (1977). Transfer of managers as a coordination and control strategy in multinational organizations. *Administrative Science Quarterly*, pages 248–263.
- Efron, B. and Tibshirani, R. (1986). Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy. *Statistical Science*, pages 54–75.
- Ertug, G., Gargiulo, M., Galunic, C., and Zou, T. (2018). Homophily and individual performance. *Organization Science*, 29(5):912–930.
- Fang, D., Moy, E., Colburn, L., and Hurley, J. (2000). Racial and ethnic disparities in faculty promotion in academic medicine. *Jama*, 284(9):1085–1092.
- Fang, L. H. and Huang, S. (2017). Gender and connections among wall street analysts. *The Review of Financial Studies*, 30(9):3305–3335.
- Felisberti, F. M. and Sear, R. (2014). Postdoctoral researchers in the UK: a snapshot at factors affecting their research output. *PloS One*, 9(4).
- Fernandez, R. M., Castilla, E. J., and Moore, P. (2000). Social capital at work: Networks and employment at a phone center. *American Journal of Sociology*, 105(5):1288–1356.
- Finn, M. G. (2011). Stay rates of foreign doctorate recipients from US universities, 2011. *Oak Ridge, TN: Oak Ridge Institute for Science and Education*.
- Fleming, L. (2001). Recombinant uncertainty in technological search. *Management Science*, 47(1):117–132.

- 
- Fox, M. F. (1983). Publication productivity among scientists: A critical review. *Social Studies of Science*, 13(2):285–305.
- Fox, M. F. and Faver, C. A. (1985). Men, women, and publication productivity: Patterns among social work academics. *The Sociological Quarterly*, 26(4):537–549.
- Fox, M. F. and Stephan, P. E. (2001). Careers of young scientists: Preferences, prospects and realities by gender and field. *Social Studies of Science*, 31(1):109–122.
- Franco, C. P. and McCowan, T. (2020). Rewiring higher education for the sustainable development goals: The case of the intercultural university of Veracruz, Mexico. *Higher Education*, pages 1–20.
- Frank, R. H. (1985). *Choosing the right pond: Human behavior and the quest for status*. Oxford University Press.
- Frink, D. D., Robinson, R. K., Reithel, B., Arthur, M. M., Ammeter, A. P., Ferris, G. R., Kaplan, D. M., and Morrisette, H. S. (2003). Gender demography and organization performance: A two-study investigation with convergence. *Group & Organization Management*, 28(1):127–147.
- Gaule, P. and Piacentini, M. (2018). An advisor like me? Advisor gender and post-graduate careers in Science. *Research Policy*, 47(4):805–813.
- Gay, V. (2019). The legacy of the missing men: The long-run impact of World War I on female labor force participation. *Available at SSRN 3069582*.
- Gemici, A. and Wiswall, M. (2014). Evolution of gender differences in post-secondary human capital investments: College majors. *International Economic Review*, 55(1):23–56.
- Gerhards, J., Hans, S., and Drewski, D. (2018). Global inequality in the academic system: effects of national and university symbolic capital on international academic mobility. *Higher Education*, 76(4):669–685.
- Gersick, C. J., Dutton, J. E., and Bartunek, J. M. (2000). Learning from academia: The importance of relationships in professional life. *Academy of Management Journal*, 43(6):1026–1044.
- Giorgi, G. (2012). Workplace bullying in academia creates a negative work environment. An Italian study. *Employee Responsibilities and Rights Journal*, 24(4):261–275.
- Girves, J. E., Zepeda, Y., and Gwathmey, J. K. (2005). Mentoring in a post-affirmative action world. *Journal of Social Issues*, 61(3):449–479.

- 
- Golde, C. M. (2005). The role of the department and discipline in doctoral student attrition: Lessons from four departments. *The Journal of Higher Education*, 76(6):669–700.
- Goldin, C. and Olivetti, C. (2013). Shocking labor supply: A reassessment of the role of World War II on women’s labor supply. *American Economic Review*, 103(3):257–62.
- Gompers, P. A. and Wang, S. Q. (2017). And the children shall lead: Gender diversity and performance in venture capital. Technical report, National Bureau of Economic Research.
- Gonzalez-Sauri, M. and Rossello, G. (2019). The role of early-career university prestige stratification on the future academic performance of scholars. *UNU-MERIT Working Paper Series*, 2019(018).
- Granovetter, M. S. (1973). The strength of weak ties. *American Journal of Sociology*, 78(6):1360–1380.
- Hagstrom, W. O. (1971). Inputs, outputs, and the prestige of university Science departments. *Sociology of Education*, pages 375–397.
- Han, J. and Han, J. (2009). Network-based recruiting and applicant attraction in China: Insights from both organizational and individual perspectives. *The International Journal of Human Resource Management*, 20(11):2228–2249.
- Hart, R. A. (2009). Did british women achieve long-term economic benefits from working in essential WWII industries?
- Hayter, C. S. and Parker, M. A. (2019). Factors that influence the transition of university postdocs to non-academic scientific careers: An exploratory study. *Research Policy*, 48(3):556–570.
- Headworth, S. and Freese, J. (2016). Credential privilege or cumulative advantage? Prestige, productivity, and placement in the academic sociology job market. *Social Forces*, 94(3):1257–1282.
- Helmer, M., Schottdorf, M., Neef, A., and Battaglia, D. (2017). Gender bias in scholarly peer review. *Elife*, 6:e21718.
- Henning, M. A., Zhou, C., Adams, P., Moir, F., Hobson, J., Hallett, C., and Webster, C. S. (2017). Workplace harassment among staff in higher education: A systematic review. *Asia Pacific Education Review*, 18(4):521–539.
- Herman, C. (2017). Looking back at doctoral education in South Africa. *Studies in Higher Education*, 42(8):1437–1454.

- 
- Hewstone, M. and Swart, H. (2011). Fifty-odd years of inter-group contact: From hypothesis to integrated theory. *British Journal of Social Psychology*, 50(3):374–386.
- Hezlett, S. A. and Gibson, S. K. (2007). Linking mentoring and social capital: Implications for career and organization development. *Advances in Developing Human Resources*, 9(3):384–411.
- Holman, L., Stuart-Fox, D., and Hauser, C. E. (2018). The gender gap in Science: How long until women are equally represented? *PLoS Biology*, 16(4):e2004956.
- Horta, H., Cattaneo, M., and Meoli, M. (2018). PhD funding as a determinant of PhD and career research performance. *Studies in Higher Education*, 43(3):542–570.
- Hottenrott, H. and Lawson, C. (2017). Flying the nest: How the home department shapes researchers career paths. *Studies in Higher Education*, 42(6):1091–1109.
- Hsieh, C.-T., Hurst, E., Jones, C. I., and Klenow, P. J. (2013). The allocation of talent and U.S. economic growth. Technical report, National Bureau of Economic Research.
- Hunt, J. (2016). Why do women leave Science and engineering? *ILR Review*, 69(1):199–226.
- Hurlbert, J. S. and Rosenfeld, R. A. (1992). Getting a good job: Rank and institutional prestige in academic psychologists' careers. *Sociology of Education*, pages 188–207.
- Ibarra, H. (1992). Homophily and differential returns: Sex differences in network structure and access in an advertising firm. *Administrative Science Quarterly*, pages 422–447.
- Ibarra, H. (1997). Paving an alternative route: Gender differences in managerial networks. *Social Psychology Quarterly*, pages 91–102.
- Jackson, M. O. (2005). A survey of network formation models: Stability and efficiency. *Group Formation in Economics: Networks, Clubs, and Coalitions*, pages 11–49.
- Jackson, M. O. (2010). *Social and economic networks*. Princeton University Press.
- Jacobi, M. (1991). Mentoring and undergraduate academic success: A literature review. *Review of Educational Research*, 61(4):505–532.
- Jadidi, M., Karimi, F., Lietz, H., and Wagner, C. (2018). Gender disparities in Science? Dropout, productivity, collaborations and success of male and female computer scientists. *Advances in Complex Systems*, 21(03n04):1750011.

- 
- Johny, J., Wichmann, B., and Swallow, B. M. (2017). Characterizing social networks and their effects on income diversification in rural Kerala, India. *World Development*, 94:375–392.
- Jost, J. T., Banaji, M. R., and Nosek, B. A. (2004). A decade of system justification theory: Accumulated evidence of conscious and unconscious bolstering of the status quo. *Political Psychology*, 25(6):881–919.
- Judge, T. A., Kammayer-Mueller, J., and Bretz, R. D. (2004). A longitudinal model of sponsorship and career success: A study of industrial-organizational psychologists. *Personnel Psychology*, 57(2):271–303.
- Jung, J. and Lee, S. J. (2016). Influence of university prestige on graduate wage and job satisfaction: The case of South Korea. *Journal of Higher Education Policy and Management*, 38(3):297–315.
- Kalaitzidakis, P., Mamuneas, T. P., and Stengos, T. (2003). Rankings of academic journals and institutions in economics. *Journal of the European Economic Association*, 1(6):1346–1366.
- Keashly, L. and Neuman, J. H. (2010). Faculty experiences with bullying in higher education: Causes, consequences, and management. *Administrative Theory & Praxis*, 32(1):48–70.
- Kelchtermans, S. and Veugelers, R. (2011). The great divide in scientific productivity: Why the average scientist does not exist. *Industrial and Corporate Change*, 20(1):295–336.
- Kilduff, M. and Lee, J. W. (2020). The integration of people and networks. *Annual Review of Organizational Psychology and Organizational Behavior*, 7.
- Koberg, C. S., Boss, R. W., and Goodman, E. (1998). Factors and outcomes associated with mentoring among healthcare professionals. *Journal of Vocational Behavior*, 53(1):58–72.
- Kolev, J., Fuentes-Medel, Y., and Murray, F. (2019). Is blinded review enough? How gendered outcomes arise under anonymous evaluation. In *Academy of Management Proceedings*, volume 2019, page 15210. Academy of Management Briarcliff Manor, NY 10510.
- Kossinets, G. and Watts, D. J. (2009). Origins of homophily in an evolving social network. *American Journal of Sociology*, 115(2):405–450.
- Kuhn, T. S. (2012). *The structure of scientific revolutions*. University of Chicago press.
- Kullback, S. and Leibler, R. A. (1951). On information and sufficiency. *The Annals of Mathematical Statistics*, 22(1):79–86.

- 
- Lang, L., Wang, Y., Chen, Q., and Zheng, T. (2019). University ranking based on faculty hiring network with minimum weighted violation rankings. *International Journal of Modern Physics C*, 30(07):1940017.
- Larivière, V. (2012). On the shoulders of students? the contribution of PhD students to the advancement of knowledge. *Scientometrics*, 90(2):463–481.
- Laurance, W. F., Useche, D. C., Laurance, S. G., and Bradshaw, C. J. (2013). Predicting publication success for biologists. *BioScience*, 63(10):817–823.
- Lawson, C. and Shibayama, S. (2015). Appointment, promotion, and mobility of bioscience researchers in Japan. In *Global Mobility of Research Scientists*, pages 239–269. Elsevier.
- Lee, D. H. (2019). Predicting the research performance of early career scientists. *Scientometrics*, 121(3):1481–1504.
- Leflot, G., Onghena, P., and Colpin, H. (2010). Teacher–child interactions: relations with children’s self-concept in second grade. *Infant and Child Development*, 19(4):385–405.
- Lehohla, P. (2015). Census 2011: Fertility in South Africa. *Statistics South Africa*. <http://www.statssa.gov.za/publications/Report-03-01-63/Report-03-01-632011.pdf>. Accessed September, 27:2018.
- León, L. R., Cowan, R., and Müller, M. (2016). Formation and persistence of research communities in Middle Income Countries: The case of South Africa. *UNU-MERIT Working Paper*.
- Lerchenmueller, M. J. and Sorenson, O. (2018). The gender gap in early career transitions in the life sciences. *Research Policy*, 47(6):1007–1017.
- Levine, R. A. and Campbell, D. T. (1972). Ethnocentrism: Theories of conflict, ethnic attitudes, and group behavior.
- Levy, H. (1992). Stochastic dominance and expected utility: Survey and analysis. *Management Science*, 38(4):555–593.
- Lewellyn, K. B. and Muller-Kahle, M. I. (2019). The corporate board glass ceiling: The role of empowerment and culture in shaping board gender diversity. *Journal of Business Ethics*, pages 1–18.
- Li, R., Zhang, H., Zhu, X., and Li, N. (2020). Does employee humility foster performance and promotability? Exploring the mechanisms of LMX and peer network centrality in China. *Human Resource Management*.
- Li, W., Aste, T., Caccioli, F., and Livan, G. (2019). Early coauthorship with top scientists predicts success in academic careers. *Nature Communications*, 10(1):1–9.



- 
- Lockwood, P. (2006). Someone like me can be successful: Do college students need same-gender role models? *Psychology of Women Quarterly*, 30(1):36–46.
- Long, J. S. (1978). Productivity and academic position in the scientific career. *American Sociological Review*, pages 889–908.
- Long, J. S., Allison, P. D., and McGinnis, R. (1979). Entrance into the academic career. *American Sociological Review*, 44(5):816–830.
- Ma, Y., Mukherjee, S., and Uzzi, B. (2020). Mentorship and protégé success in STEM fields. *Proceedings of the National Academy of Sciences*.
- Machado, J., Parente, P., and Silva, J. S. (2011). Qreg2: Stata module to perform quantile regression with robust and clustered standard errors. Statistical Software Components, Boston College Department of Economics.
- Mai, B., Liu, J., and González-Bailón, S. (2015). Network effects in the academic market: Mechanisms for hiring and placing PhDs in communication (2007-2014). *Journal of Communication*, 65(3):558–583.
- Main, J. B. (2014). Gender homophily, Ph.D. completion, and time to degree in the humanities and humanistic Social Sciences. *The Review of Higher Education*, 37(3):349–375.
- Mairesse, J. and Pezzoni, M. (2015). Does gender affect scientific productivity? *Revue Économique*, 66(1):65–113.
- Mairesse, J., Pezzoni, M., and Visentin, F. (2019). Impact of family characteristics on the gender publication gap: Evidence for physicists in France. *Interdisciplinary Science Reviews*, 44(2):204–220.
- Marsh, H. W. (1991). Failure of high-ability high schools to deliver academic benefits commensurate with their students ability levels. *American Educational Research Journal*, 28(2):445–480.
- Marsh, H. W. and Hau, K.-T. (2003). Big-fish-little-pond effect on academic self-concept: A cross-cultural (26-country) test of the negative effects of academically selective schools. *American Psychologist*, 58(5):364.
- Marsh, H. W., Seaton, M., Trautwein, U., Lüdtke, O., Hau, K.-T., O Mara, A. J., and Craven, R. G. (2008). The big-fish-little-pond-effect stands up to critical scrutiny: Implications for theory, methodology, and future research. *Educational Psychology Review*, 20(3):319–350.
- Mas-Colell, A., Whinston, M. D., Green, J. R., et al. (1995). *Microeconomic theory*, volume 1. Oxford University Press New York.

- 
- McKay, R., Arnold, D. H., Fratzl, J., and Thomas, R. (2008). Workplace bullying in academia: A Canadian study. *Employee Responsibilities and Rights Journal*, 20(2):77–100.
- McKeown, S. and Dixon, J. (2017). The contact hypothesis: Critical reflections and future directions. *Social and Personality Psychology Compass*, 11(1):e12295.
- McPherson, J. M. and Smith-Lovin, L. (1987). Homophily in voluntary organizations: Status distance and the composition of face-to-face groups. *American Sociological Review*, pages 370–379.
- McPherson, M., Smith-Lovin, L., and Cook, J. M. (2001). Birds of a feather: Homophily in social networks. *Annual Review of Sociology*, 27(1):415–444.
- Merton, R. K. (1968). The Matthew effect in Science. *Science*, 159(3810):56–63.
- Merton, R. K. (1988). The Matthew effect in Science, II: Cumulative advantage and the symbolism of intellectual property. *Isis*, 79(4):606–623.
- Milkman, K. L., Akinola, M., and Chugh, D. (2015). What happens before? A field experiment exploring how pay and representation differentially shape bias on the pathway into organizations. *Journal of Applied Psychology*, 100(6):1678.
- Miller, C. C., Glick, W. H., and Cardinal, L. B. (2005). The allocation of prestigious positions in organizational Science: Accumulative advantage, sponsored mobility, and contest mobility. *Journal of Organizational Behavior*, 26(5):489–516.
- Mollica, K. A., Gray, B., and Treviño, L. K. (2003). Racial homophily and its persistence in newcomers social networks. *Organization Science*, 14(2):123–136.
- Montgomery, J. D. (1991). Social networks and labor-market outcomes: Toward an economic analysis. *The American Economic Review*, 81(5):1408–1418.
- Moodie, G. (2009). Four tiers. *Higher Education*, 58(3):307–320.
- Moore, K. K., Cid-Martinez, I., Toney, J., Smith, J. A., Kalb, A. C., Shin, J. H., and Spalter-Roth, R. M. (2018). Who climbs the academic ladder? Race and gender stratification in a world of whiteness. *The Review of Black Political Economy*, 45(3):216–244.
- Moss-Racusin, C. A., van der Toorn, J., Dovidio, J. F., Brescoll, V. L., Graham, M. J., and Handelsman, J. (2016). A “scientific diversity” intervention to reduce gender bias in a sample of life scientists. *CBE-LIFE SCIENCES EDUCATION*, 15(3).
- Mouton, Van Lill, B., Boshoff, V., Cloete, and Sheppard (2015). A study on the retention, completion and progression rates of South African postgraduate students. *Research Report*.

- 
- Mullen, A. L. and Baker, J. (2008). Gender, race, and ethnic segregation of Science fields in US universities. *Journal of Women and Minorities in Science and Engineering*, 14(2).
- Nchinda, T. C. (2002). Research capacity strengthening in the South. *Social Science & Medicine*, 54(11):1699–1711.
- Nevin, A. D. (2019). Academic hiring networks and institutional prestige: A case study of Canadian sociology. *Canadian Review of Sociology*, 56(3):389–420.
- Newman, M. E. (2003). Mixing patterns in networks. *Physical Review E*, 67(2):026126.
- Olcay, G. A. and Bulu, M. (2017). Is measuring the knowledge creation of universities possible?: A review of university rankings. *Technological Forecasting and Social Change*, 123:153–160.
- Oyarzún, J. d. D., Perales Franco, C., and McCowan, T. (2017). Indigenous higher education in Mexico and Brazil: Between redistribution and recognition. *Compare: A Journal of Comparative and International Education*, 47(6):852–871.
- Oyer, P. (2008). Ability and employer learning: Evidence from the economist labor market. *Journal of the Japanese and International Economies*, 22(2):268–289.
- Petscher, Y. and Logan, J. A. (2014). Quantile regression in the study of developmental sciences. *Child Development*, 85(3):861–881.
- Pettigrew, T. F. and Tropp, L. R. (2006). A meta-analytic test of intergroup contact theory. *Journal of Personality and Social Psychology*, 90(5):751.
- Pezzoni, M., Mairesse, J., Stephan, P., and Lane, J. (2016). Gender and the publication output of graduate students: A case study. *PLoS One*, 11(1):e0145146.
- Phelps, E. S. (1972). The statistical theory of racism and sexism. *The American Economic Review*, pages 659–661.
- Pifer, M. J. and Baker, V. L. (2013). Managing the process: The intradepartmental networks of early-career academics. *Innovative Higher Education*, 38(4):323–337.
- Pillai, K. G., Leonidou, C., and Bian, X. (2018). Let’s be friends: National homophily in multicultural newcomer student networks. *Social Networking*, 8(01):16–38.
- Pin, P. and Rogers, B. W. (2016). Stochastic network formation and homophily.
- Pinheiro, D. L., Melkers, J., and Newton, S. (2017). Take me where I want to go: Institutional prestige, advisor sponsorship, and academic career placement preferences. *PloS One*, 12(5):e0176977.
- Podolny, J. M. (1994). Market uncertainty and the social character of economic exchange. *Administrative Science Quarterly*, pages 458–483.

- 
- Ramasco, J. J. and Morris, S. A. (2006). Social inertia in collaboration networks. *Physical Review E*, 73(1):016122.
- Reskin, B. F. (1977). Scientific productivity and the reward structure of Science. *American Sociological Review*, pages 491–504.
- Reyes Ruiz, G. and Suriñach, J. (2015). Análisis sobre la evolución del sistema nacional de investigadores (SNI) de México. *Investigación Administrativa*, 44(115):55–69.
- Rivera León, L., Mairesse, J., and Cowan, R. (2017). Gender gaps and scientific productivity in middle-income countries: Evidence from Mexico. Technical report, Inter-American Development Bank.
- Rosman, T., Mayer, A.-K., Leichner, N., and Krampen, G. (2020). Putting big fish into a bigger pond: self-concept changes in psychology undergraduate entrants. *Journal of Further and Higher Education*, 44(1):14–28.
- Rossello, G. and Cowan, R. (2019). Far from random? the role of homophily in student supervision. *UNU-MERIT Working Paper Series*, 2019(024).
- Rossello, G., Cowan, R., and Mairesse, J. (2020). Ph. D. research output in STEM: the role of gender and race in supervision. *UNU-MERIT Working Paper Series*, 2020(021).
- Rossiter, M. W. (1993). The Matthew Matilda effect in Science. *Social Studies of Science*, 23(2):325–341.
- Rubie-Davies, C. M. (2006). Teacher expectations and student self-perceptions: Exploring relationships. *Psychology in the Schools*, 43(5):537–552.
- Rzhetsky, A., Foster, J. G., Foster, I. T., and Evans, J. A. (2015). Choosing experiments to accelerate collective discovery. *Proceedings of the National Academy of Sciences*, 112(47):14569–14574.
- Salchegger, S. (2016). Selective school systems and academic self-concept: How explicit and implicit school-level tracking relate to the big-fish-little-pond effect across cultures. *Journal of Educational Psychology*, 108(3):405.
- Saleh, A., Yu, Q., Leslie, S. H., and Seydel, J. (2017). Gender equity, student loans and returns on investment in American higher education. *International Journal of Sociology of Education*, 6(2):216–243.
- Sarsons, H. et al. (2015). Gender differences in recognition for group work. *Harvard University*, 3.
- Schmutte, I. (2016). How do social networks affect labor markets? *IZA World of Labor*.

- 
- Sekhon, J. S. (2011). Multivariate and propensity score matching software with automated balance optimization: The Matching package for R. *Journal of Statistical Software*, 42(7):1–52.
- Sharmeen, F., Arentze, T., and Timmermans, H. (2014). Dynamics of face-to-face social interaction frequency: Role of accessibility, urbanization, changes in geographical distance and path dependence. *Journal of Transport Geography*, 34:211–220.
- Shavit, Y. (2007). *Stratification in higher education: A comparative study*. Stanford University Press.
- Sherf, E. N. and Morrison, E. W. (2019). I do not need feedback! Or do I? self-efficacy, perspective taking, and feedback seeking. *Journal of Applied Psychology*.
- Sherif, M. (1966). *Group conflict and co-operation: Their social psychology*. London: Routledge and Kegan Paul.
- Sherif, M. (2015). *Group conflict and co-operation: Their social psychology*. Psychology Press.
- Sherif, M., Harvey, O., Hood, W., and Sherif, C. (1961). Intergroup conflict and group relations: The robbers cave experiment. *The Institute of Group Relations, Norman, OK*.
- Shi, F., Foster, J. G., and Evans, J. A. (2015). Weaving the fabric of Science: Dynamic network models of Science’s unfolding structure. *Social Networks*, 43:73–85.
- Skvoretz, J. (2013). Diversity, integration, and social ties: Attraction versus repulsion as drivers of intra- and intergroup relations. *American Journal of Sociology*, 119(2):486–517.
- Slavova, K., Fosfuri, A., and De Castro, J. O. (2016). Learning by hiring: The effects of scientists inbound mobility on research performance in academia. *Organization Science*, 27(1):72–89.
- Smith, J. A., McPherson, M., and Smith-Lovin, L. (2014). Social distance in the United States: Sex, race, religion, age, and education homophily among confidants, 1985 to 2004. *American Sociological Review*, 79(3):432–456.
- Somers, M. A., Cabus, S. J., Groot, W., and van den Brink, H. M. (2019). Horizontal mismatch between employment and field of education: Evidence from a systematic literature review. *Journal of Economic Surveys*, 33(2):567–603.
- Stainback, K. (2008). Social contacts and race/ethnic job matching. *Social Forces*, 87(2):857–886.
- Stapel, D. A. and Koomen, W. (2005). Competition, cooperation, and the effects of others on me. *Journal of personality and social psychology*, 88(6):1029.

- 
- Steele, J. R. and Ambady, N. (2006). Math is hard! the effect of gender priming on women's attitudes. *Journal of Experimental Social Psychology*, 42(4):428–436.
- Stephan, P. and Ma, J. (2005). The increased frequency and duration of the postdoctorate career stage. *American Economic Review*, 95(2):71–75.
- Su, X. (2011). Postdoctoral training, departmental prestige and scientists' research productivity. *The Journal of Technology Transfer*, 36(3):275–291.
- Teplitskiy, M., Duede, E., Menietti, M., and Lakhani, K. R. (2020). Citations systematically misrepresent the quality and impact of research articles: Survey and experimental evidence from thousands of citers. *arXiv preprint arXiv:2002.10033*.
- Terrell, M. C. and Hassell, R. K. (1994). Mentoring undergraduate minority students: An overview, survey, and model program. *New Directions for Teaching and Learning*, 1994(57):35–45.
- Tiedemann, J. (2000). Parents' gender stereotypes and teachers' beliefs as predictors of children's concept of their mathematical ability in elementary school. *Journal of Educational Psychology*, 92(1):144.
- Uzzi, B., Mukherjee, S., Stringer, M., and Jones, B. (2013). Atypical combinations and scientific impact. *Science*, 342(6157):468–472.
- Van den Brink, M. and Benschop, Y. (2014). Gender in academic networking: The role of gatekeepers in professorial recruitment. *Journal of Management Studies*, 51(3):460–492.
- Van Der Lee, R. and Ellemers, N. (2015). Gender contributes to personal research funding success in the Netherlands. *Proceedings of the National Academy of Sciences*, 112(40):12349–12353.
- Verdín, D., Godwin, A., Kirn, A., Benson, L., and Potvin, G. (2018). Engineering women's attitudes and goals in choosing disciplines with above and below average female representation. *Social Sciences*, 7(3):44.
- Verhoeven, D., Bakker, J., and Veugelers, R. (2016). Measuring technological novelty with patent-based indicators. *Research Policy*, 45(3):707–723.
- Von Keyserlingk, L., Becker, M., and Jansen, M. (2020). Do social comparisons matter for university major choices? a longitudinal study from a gender perspective. *International Journal of Gender, Science and Technology*, 12(1):46–64.
- Vries, H. d. (1998). Finding a dominance order most consistent with a linear hierarchy: A new procedure and review. *Animal Behaviour*, 55(4):827–843.
- Wager, E., Singhvi, S., and Kleinert, S. (2015). Too much of a good thing? An observational study of prolific authors. *PeerJ*, 3:e1154.

- 
- Wang, J., Veugelers, R., and Stephan, P. (2017). Bias against novelty in Science: A cautionary tale for users of bibliometric indicators. *Research Policy*, 46(8):1416–1436.
- Wang, Y., Jones, B. F., and Wang, D. (2019). Early-career setback and future career impact. *Nature Communications*, 10(1):1–10.
- Watson, J. (2012). Networking: Gender differences and the association with firm performance. *International Small Business Journal*, 30(5):536–558.
- Webster, D. S., Conrad, C. F., and Jensen, E. L. (1991). Objective and reputational rankings of Ph.D. granting departments of sociology, 1965–1982. *Sociological Focus*, 21(2):177–198.
- Weeden, K. A., Thébaud, S., and Gelbgiser, D. (2017). Degrees of difference: Gender segregation of US doctorates by field and program prestige. *Sociological Science*, 4:123–150.
- Weiss, H. M. and Knight, P. A. (1980). The utility of humility: Self-esteem, information search, and problem-solving efficiency. *Organizational Behavior and Human Performance*, 25(2):216–223.
- Wenger, E. (2010). Communities of practice and social learning systems: The career of a concept. In *Social learning systems and communities of practice*, pages 179–198. Springer.
- Werts, C. E. and Watley, D. J. (1969). A student's dilemma: Big fish-little pond or little fish-big pond. *Journal of Counseling Psychology*, 16(1):14.
- Westphal, J. D. and Zhu, D. H. (2019). Under the radar: How firms manage competitive uncertainty by appointing friends of other chief executive officers to their boards. *Strategic Management Journal*, 40(1):79–107.
- Whelpley, C. E. and McDaniel, M. A. (2016). Self-esteem and counterproductive work behaviors: A systematic review. *Journal of Managerial Psychology*.
- Williamson, I. O. and Cable, D. M. (2003). Predicting early career research productivity: The case of management faculty. *Journal of Organizational Behavior*, 24(1):25–44.
- Zebrowitz, L. A., White, B., and Wieneke, K. (2008). Mere exposure and racial prejudice: Exposure to other-race faces increases liking for strangers of that race. *Social Cognition*, 26(3):259–275.
- Zeltzer, D. (2019). Gender homophily in referral networks: Consequences for the medicare physician earnings gap. *Available at SSRN 2921482*.

- 
- Zhou, S., Page-Gould, E., Aron, A., Moyer, A., and Hewstone, M. (2018). The extended contact hypothesis: A meta-analysis on 20 years of research. *Personality and Social Psychology Review*.
- Zhu, Y. and Yan, E. (2017). Examining academic ranking and inequality in library and information Science through faculty hiring networks. *Journal of Informetrics*, 11(2):641–654.
- Zinovyeva, N. and Bagues, M. (2010). Does gender matter for academic promotion? evidence from a randomized natural experiment.
- Zinovyeva, N. and Bagues, M. (2015). The role of connections in academic promotions. *American Economic Journal: Applied Economics*, 7(2):264–92.
- Zuckerman, H. (1977). *Scientific elite: Nobel laureates in the United States*. Transaction Publishers.



---

# Valorisation Addendum

In line with Article 23 of the regulation governing the attainment of the doctoral degree at Maastricht University, in this section, I highlight the valorisation opportunities of this doctoral dissertation. The aim of this section is to show how the knowledge developed in my research project can be valuable for society, policy makers, and other scholars.

The general inspiration of this dissertation was to examine why specific demographic groups remain under-represented in knowledge-intensive occupations. I examined this with the empirical focus of higher education. The standard economic theory predicts that rational agents will recruit workers based on individual merit, skills, and intelligence and not based on physical appearance or group belonging. The Economic theory considers demographic imbalances as sub-optimal and destined to disappear. However, the Nobel laureate Gary Becker highlighted that this might not be the case when agents have a taste for discrimination. Recent contributions in sociology and psychology examined how this consideration might be more nuanced, showing that this “taste for discrimination” might be unconscious. Moreover, in educational settings, merit and skills of students are endogenous to the opportunities and the quality of supervision is provided them. Since humans tend to display an homophilistic tendency, under-represented groups are disadvantaged because they have fewer opportunities to meet similar ones. This observation is the heart of my work and implies that individual and systemic mechanisms together play a role in perpetuating these demographic imbalances.

In addition to the scientific relevance discussed in the previous chapters, this dissertation contains research results related to this issue that are of socio-economic relevance for our society. Most of them were highlighted in the previous chapters, I summarise them in this valorisation addendum.

Beyond the general framework, each chapter examined systemic and individual mechanisms of the entry process in the workplace. The chapters of this thesis provided interesting results and new empirical methods.

In Chapter 2, I found that the formation of student-supervisor relation can operate as a first bottleneck preventing social transformations.

---

I developed a new methodology to separate choice homophily from that induced by the system which solves causal issues creating a null model of type-blind tie formation that incorporates network structures and individual characteristics (i.e. supervision capacity, popularity, department norms). Using those results, I examined how population groups relate to one other. I developed a model with interacting populations and estimated homophilous preferences in the demographic groups using aggregate department level data of student, supervisor, and tie composition. The methodology can be used for monitoring social transformations within organisations and it is suitable in the presence of sparse register data. I applied the described methods to register data of student-advisor relations at South African universities between 1973 and 2014.

My results underline that in the university system both race-based and gender-based homophily influence tie formation. The observed racial (gender) homophily in the system arises from institutional constraints, network structures, and choices made by individuals. In particular, half of the observed homophily is induced by the system through institutional constraints and network structures and half is driven by choices. The latter originates primarily in new members of the former dominant group (white/male students). However, when the sample of students is restricted to promising students who enter academia, white professors also display choice homophily. Further, I found that during a transition phase of social transformations homophily can mechanically increase even if there are no behavioural changes among individuals, under certain circumstances. In terms of policy recommendations, Chapter 2 underlines that to achieve sustainable social transformations a demographic change is not enough and that structural constraints and behavioural aspects should be tackled together. In particular, white (male) students and their advisors is a critical target group for intervention.

In Chapter 3, I focussed on the individual mechanisms related to training, testing the link between homophilous/heterophilous (same-/cross- type) ties and productivity. In particular, I studied whether student-advisor gender and racial couple affects student's doctoral productivity. I used an econometric technique, employing a panel data analysis with a quantile regression approach. This approach studies the link between homophilous/heterophilous ties and productivity looking at different student productivity-profiles. I found that on average female students publish less than males. This difference exists for female students working with a male advisor (heterophilous ties) and not those working with a female supervisor (homophilous ties). But looking at this link for different student productivity-profiles I found the following. In the whole sample, female students with a high (or low) productivity-profile studying with female advisors are as productive as male students with a high (or low) productivity-profile studying with male advisors. Decomposing the joint effect of gender and race I found different results looking at same- and cross-racial supervision. For same-race supervision, the gap in average performance between female students and male students is mainly driven by a gap in the right hand tails of the productivity distribution (students with high productivity profiles). In contrast, cross-racial supervisions display a smaller gender gap that is u-shaped over student productivity.

---

Chapter 3 underlines the role that female advisors have for female students. Further, it shows the crucial role that cross-racial ties have in reducing gender productivity differences. Moreover, the chapter points out that same-(cross-) racial ties have a heterogeneous effect on student productivity depending on her/his productivity-profile. In terms of policy advice, the evidence from the chapter shows that promoting cross-racial ties and female supervision have potential in decreasing the gender gap.

Most of us agree that achieving a balanced representation of the population in desirable occupations is politically and socially important. Fulfilling a better balance can enhance economic growth and social development because it increases the pool of talents to select from, and empowers those previously excluded. However, Chapters 2 and 3 highlighted how the inclusion of previously discriminated groups could be more difficult than expected and hard to implement with a single policy target. More importantly, the main results of Chapters 2 and 3 challenge the current policy approach to integration and inclusion that considers the under-represented population as the main target group. On the one hand, Chapters 2 and 3 suggest that young agents of the former majority groups (white males) are should be the target of policies. On the other hand, Chapter 3 highlights that the main target should be the group of highly productive students where most of the student gender productivity gap is generated. The shift in integration and inclusion policy design, suggested by my findings in Chapters 2 and 3, has a direct implication to strategies aimed at achieving the following UN Sustainable Development Goals (SDG): UN-SDG number 4 Quality Education; UN-SDG number 5 Gender Equality; and UN-SDG number 10 Reduced Inequalities. Policies devoted to promoting such goals should have a holistic and long-run perspective beyond labour markets outcomes, and crucially they must focus on the group of people that hardly conceives possible gains from such policies.

The second part of the dissertation, in Chapters 4 and 5, I examined university rankings and their implication in the Ph.D. job market and future performance. These Chapters connect with the increasing importance that universities play in our society, and the growing interest in the evaluation and ranking of universities. In particular, university rankings are often used in the job market to screen job candidates. However, most of university rankings have measuring issues, and they mainly lack information about how peers and competitor academics evaluate each other. To solve this issue, Chapters 4 and 5 develop a new ranking algorithm of universities bases on hiring networks. This contributes to the policy debate since an improved measure of university prestige permits a better evaluation of universities for those considering career. In particular, I studied the link between hiring, university prestige, and research quality. I developed a technique to rank university prestige and used it to test whether prestige movements affect scholar research quality. To achieve this, I developed a new quasi-experimental method, based on matching. I compared the quality performance of matched pairs of individuals with the same training (or same job) but different movements in the prestige hierarchy.

I found that the university system has a hierarchical prestige ordering with specific

---

hiring flows. In this stratified hierarchy, blacks and females are disadvantaged — they tend to experience more mobility and to move down in the prestige hierarchy when hired in their first job in academia. Overall, inertia has a positive role in the system: those who make large movements in terms of prestige have lower research ratings than those who are hired by the same university granting the Ph.D.. This inertia also affects the long run (20 years after Ph.D. graduation). The results in the Chapters suggest that the hierarchical stratification of the university system is a constraint to social transformation because it systematically disadvantages upward prestige mobility of black and female Ph.D.s, in particular, with effects also in the long-run. In terms of policy, any intervention aimed at counteracting this hierarchical stratification allowing a fruitful exchange between institutions will promote social transformations. Further, developing incentives to allow black and female students to stay after Ph.D. and build their research capacity before changing institution could also be beneficial.

Chapters 4 and 5 are of particular interests for policy makers interested in the diffusion of knowledge and talents. My findings challenge the common notion that mobility positively affects the performance of a researcher. The association between mobility and performance is non-linear and context-specific. Indeed, most of the results that link mobility and performance are from North America where there are no language differences, and where mobility is at the base of cultural norms (see for example the “American Dream”). Moreover, North America and the US, in particular, are attractors of talents meaning that also the process of including a new researcher member joining the team is common. A similar organisational and routine activity might differ in a context where mobility is rare.

The research results of my dissertation are therefore of vital interest for policy makers and university rectors that want to achieve effective inclusion policies, to increase productivity, and to create a better research environment.

In my thesis, I developed new methods. Chapter 2 develops a new methodology which separates choice homophily from that induced by the system. The method solves causal issues creating a null model of random and unbiased tie formation that controls for network structures, supervision capacity, and small sample bias. Further, I design a model with interacting populations and use it to estimate homophilous preferences using aggregate department-level data of student, supervisor, and tie composition. The methodology is promising for monitoring aggregate and group level homophily in the entry process. Further, the method can also be used with sparse register data.

Chapter 3 uses an econometric technique. In particular, I employ a panel data analysis with a quantile regression approach. This to investigate the link between student-advisor same- and cross-gender (and race) couples and productivity. In this chapter, I investigate the relation between homophily and productivity in the different student productivity-profiles. This accounts for the possibility that homophily/heterophily (same-/cross-gender -race ties) works differently for (hub) scholars with extraordinary

---

levels of productivity who experience Matthew effects.

Chapter 4 develops a new technique to study prestige stratification of institutions. Based on hiring networks, the algorithm ranks universities according to their reputation among academics. Further, I develop a new quasi-experiment method to test the effect of movements in the prestige hierarchy on research quality. The method is based on exact matching and compares the performance of matched pairs of scholars with the same training (or job) but different prestige transitions.

Chapter 5 modifies the ranking algorithm to rank university prestige. The modified version of the algorithm takes into account that universities may recruit in a cyclical manner. In this way, the university ranking is dynamic over time and accounts for changes in investment flows, modes, news. Then, I apply the matching methodology used in Chapter 4.

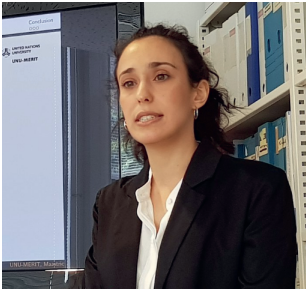
The new methodologies developed in my thesis can be of interest for concrete product developments. For example, they can be implemented into statistical packages like R or STATA and used for policy evaluation. This future development is beyond the scope of my research but is a possible commercial activity. Indeed most of my methods are suited for application in different organisations (public agencies, large firms, ...).

The innovation content of this dissertation stands in two main points. First, it is in its methodology. The various techniques I developed are tailored to the characteristics of the data I analysed and are particularly suited for sparse register data. Second, some of my results are counter-intuitive or in the literature similar results have been supported by little empirical evidence, mostly based on anecdotes.

Regarding the plans for further dissemination and valorisation of my results, I will present my work to the academic community through conferences and workshops. Hopefully, the chapters of my dissertation will be published in international peer-review journals. One has been recently published by *Scientometrics*.

---

# About the Author



**Giulia Rossello** is a PhD candidate in Economics at UNU-MERIT, Maastricht University (The Netherlands) and she is currently post-doc researcher at Sant’Anna school of advanced studies in Pisa (Italy). Her work is at the intersection of economics of knowledge and innovation, education, science of science and sociology. In particular, her research uses network analysis and theory to uncover mechanisms at work in the entry process of scientists in academia. Her empirical focus is on the emerging science systems of South Africa and Mexico. Her special interest is to understand how the stratification of university prestige, racial and gender homophily, and Matthew effects act at various level of aggregation and in relation to social transformations.

She holds a Bachelor degree in Economics from Pisa University and a Master degree in Economics from Pisa University and Sant’Anna school of advanced studies. During her doctorate, she was a visiting student at the Centre for Research on Evaluation of Science and Technology (CREST) at Stellenbosch University (South Africa). She was teaching assistant of Data Science and Econometrics in the Master of Science in Public Policy and Human Development at UNU-MERIT and of Network Economics at the School of Business and Economics at Maastricht University. Her PhD research is part of project “*Social Transformation in South African Science and Innovation*” funded by the “Agence Nationale de la Recherche” of France.

One of her PhD working papers was recently published in the journal “*Scientometrics*”. The remaining three are UNU-MERIT Working Papers and will be submitted for publication. Her research was presented at the international conferences WEHIA 2017, STI 2018, and EAEPE 2019. She presented her work as an invited seminar speaker at Università Cattolica del Sacro Cuore (Italy), Scuola Superiore Sant’Anna (Italy), and Harvard University (USA).





# UNU-MERIT/MGSoG Dissertation Series

*Change Adaptation: An Analysis of  
Adaptation Fund Projects*  
UNU-MERIT/MGSoG Dissertation  
Series № 250

**2021**

**Giulia Rossello**

*Social Transformations and Labour  
Market Entry*  
*An Investigation into University  
Systems in Emerging Economies*  
UNU-MERIT/MGSoG Dissertation  
Series № 253

**2020**

**Rose Camille Vincent**

*Essays in Public Economics*  
*Multi-Layer Tax Structure and  
Implications*  
UNU-MERIT/MGSoG Dissertation  
Series № 252

**Emmanuel Mensah**

*Structural change in developing  
countries: patterns, causes, and  
consequences*  
UNU-MERIT/MGSoG Dissertation  
Series № 251

**Ornsaran Manuamorn**

*Governance of International  
Adaptation Finance for Local Climate*

**Gillian McFarland**

*Doing policy in Further Education*  
*An exploration of the enactment of the  
GCSE resits policy in Further  
Education colleges in England*  
UNU-MERIT/MGSoG Dissertation  
Series № 249

**Omar Rodriguez Torres**

*Essays on Entrepreneurship in  
Colombia*  
UNU-MERIT/MGSoG Dissertation  
Series № 248

**Elaine Lebon-McGregor**

*International Organizations and  
Global Migration Governance*  
UNU-MERIT/MGSoG Dissertation  
Series № 247

**Janyl Moldalieva**

*Playing the 'Game' of Transparency  
and Accountability in Kyrgyzstan's  
Resource Governance*  
UNU-MERIT/MGSoG Dissertation  
Series № 246

**Kaleab Kebede Haile**

*Essays on Rural Household Decision-  
Making under Climate Risk*  
UNU-MERIT/MGSoG Dissertation  
Series № 245

**Nora Jasmin Ragab**

*Diaspora Mobilisation in a Conflict Setting*

UNU-MERIT/MGSoG Dissertation  
Series № 244

**Mary Kaltenberg**

*From Micro to Macro: Essays on Technological Change and Income Divergence*

UNU-MERIT/MGSoG Dissertation  
Series № 243

**Ayla E. Bonfiglio**

*Student migrant, refugee or both? Exploring Refugee Agency and Mobility through Tertiary Education in Kenya, South Africa and Uganda*

UNU-MERIT/MGSoG Dissertation  
Series № 242

**Danilo Sartorello Spinola**

*Cycles, Economic Structures and External Constraints.*

*A Structuralist study on the causes of economic volatility in Latin America*  
UNU-MERIT/MGSoG Dissertation  
Series № 241

**2019**

**Jemal Adem**

*Livelihood Vulnerability to Shocks, Behaviour and Investment in Education: Essays in Behavioural Development Economics*

UNU-MERIT/MGSoG Dissertation  
Series № 240

**Davina Osei**

*Corrupt or corrupted networks? An empirical enquiry*

UNU-MERIT/MGSoG Dissertation  
Series № 239

**Patima Chongcharoentanawat**

*Beyond Static Inequality Public policies and economic mobility in Thailand*

UNU-MERIT/MGSoG Dissertation  
Series № 238

**Charlotte Guillard**

*Rethinking economic growth and structural change. The role of boundaries and linkages between industries*

UNU-MERIT/MGSoG Dissertation  
Series № 237

**Nicolas Echarti**

*Employment Effects or Vocational Rehabilitation in Germany: A quantitative analysis*

UNU-MERIT/MGSoG Dissertation  
Series № 236

**Shellie E. Solomon**

*Neighborhoods Matter: Crime, collective efficacy and foreclosures in Miami*

UNU-MERIT/MGSoG Dissertation  
Series № 235

**Michał Kazimierczak**

*Regional Innovation and  
Entrepreneurship*

*Patents, trade marks, entry and  
entrants' growth in European  
manufacturing industries*  
UNU-MERIT/MGSoG Dissertation  
Series № 234

**Fernanda Soares**

*The Influence of Within School and  
Across Schools' Collaborative Practices  
on Student Learning and Teaching  
Outcomes in West Africa*  
UNU-MERIT/MGSoG Dissertation  
Series № 233

**Mira Bierbaum**

*New Mindsets to Innovate Activation*  
UNU-MERIT/MGSoG Dissertation  
Series № 232

**Norman Dytianquin**

*Technology in the Asian Miracle and  
Crisis Debates: Applications of and  
Insights from the Field of Influence  
Approach to Input-Output Analysis*  
UNU-MERIT/MGSoG Dissertation  
Series № 231

**Nga Le**

*The implications of health insurance  
for the labour market and patient  
satisfaction with medical care in  
Vietnam*  
UNU-MERIT/MGSoG Dissertation  
Series № 230

**Jinhyuck Park**

*Intellectual Property right protection  
and cross-border R&D investments by  
multinational enterprises*  
UNU-MERIT/MGSoG Dissertation  
Series № 229

**Richard de Groot**

*Show me the Money:  
Essays on the Impact of Cash Transfers  
on Child Nutrition and the Role of  
Intra-Household Dynamics*  
UNU-MERIT/MGSoG Dissertation  
Series № 228

**Catie Lott**

*Diamonds are a Women's Best Friend  
Broadening Measures of Women's  
Access to Formal Political Decision-  
Making*  
UNU-MERIT/MGSoG Dissertation  
Series № 227

**Ana Cristina Calderon Ramirez**

*Public Management Reforms  
Three stories about public procurement  
agencification in Latin America*  
UNU-MERIT/MGSoG Dissertation  
Series № 226

**Camilo Nicanor Carrillo Purin**

*Teachers' in-service training and  
student achievement:  
The effect of in-service training of  
Peruvian teachers on student  
achievement*  
UNU-MERIT/MGSoG Dissertation  
Series № 225

**Hugo Confraria**

*Developing scientific capacity in the  
Global South*

UNU-MERIT/MGSoG Dissertation  
Series № 224

**Alison Cathles**

*Educational Pathways and Skills:  
Past, Present, and Future*

UNU-MERIT/MGSoG Dissertation  
Series № 223

**Ibrahima Sory Kaba**

*Aggregate Fluctuations and  
Development: Essays on  
Macroeconomic Volatility and  
Economic Growth*

UNU-MERIT/MGSoG Dissertation  
Series № 222

**Charlotte Keijser**

*Firm Participation, Learning and  
Innovation in Heterogenous Value  
Chains of IT-enabled Services*

UNU-MERIT/MGSoG Dissertation  
Series № 221

**Salih Çevikarslan**

*Innovation Strategies and Their  
Implications for Technological Change  
and Market Outcomes:*

*An Evolutionary Multi-Agent Based  
Modelling Approach*  
UNU-MERIT/MGSoG Dissertation  
Series № 220

**Wondimagegn Mesfin Tesfaye**

*Essays on the Impacts of Climate-  
Smart Agricultural Innovations on  
Household Welfare*

UNU-MERIT/MGSoG Dissertation  
Series № 219

**Tatevik Poghosyan**

*How Board Networks Affect Firm  
Performance and Innovation Incentives  
in Transition Economies: The Case of  
Armenia*

UNU-MERIT/MGSoG Dissertation  
Series № 218

**Arip Muttaqien**

*Essays on Inequality and Polarization:  
Empirical Studies in Developing Asia*  
UNU-MERIT/MGSoG Dissertation  
Series № 217

**2018**

**Katrin Marchand**

*Essays on Forced Migration and  
Labour Market Participation in  
Developing Countries*

UNU-MERIT/MGSoG Dissertation  
Series № 216

**Ortrun Merkle**

*The Myth of Gender Neutral Power:  
Corruption and Gender Norms*  
UNU-MERIT/MGSoG Dissertation  
Series № 215

**Biljana Meshkovska**

*Life after Trafficking:  
(re)integration processes of women that  
have been trafficked for the purpose of  
sexual exploitation in Europe*  
UNU-MERIT/MGSoG Dissertation  
Series № 214

**Vincenzo Vinci**

*The Relevance of Institutions and  
People's Preferences for Social  
Protection*  
UNU-MERIT/MGSoG Dissertation  
Series № 213

**Silke Heuser**

*The Effectiveness of Environmental  
Policies on Reducing Deforestation in  
the Brazilian Amazon*  
UNU-MERIT/MGSoG Dissertation  
Series № 212

**Jennifer Waidler**

*Social Assistance and Remittances and  
Their Role in the Fight Against  
Poverty*  
UNU-MERIT/MGSoG Dissertation  
Series № 211

**Choolwe Muzyamba**

*The role of community mobilization in  
the promotion of maternal health of  
women living with HIV in Zambia*  
UNU-MERIT/MGSoG Dissertation  
Series № 210

**Juan Carlos A. Castillo Sánchez**

*Assessing the Role of the Export Sector  
in Mexican Economic  
Development, 1965-2014*  
UNU-MERIT/MGSoG Dissertation  
Series № 209

**Tareq Abuelhaj**

*Food Security Policy Impact Analysis:  
The Econometrics of Cash and Food  
Assistance Cost Effectiveness*  
UNU-MERIT/MGSoG Dissertation  
Series № 208

**Marta Fernández de Arroyabe**

**Arranz**  
*Essays on M&AS and Innovation*  
UNU-MERIT/MGSoG Dissertation  
Series № 207

**Clotilde Mahé**

*Essays on Migration and Occupational  
Choice*  
UNU-MERIT/MGSoG Dissertation  
Series № 206

**Simone Sasso**

*Talent on the move. Essays on Human  
Capital, Graduate Mobility and  
Economic Development*  
UNU-MERIT/MGSoG Dissertation  
Series № 205

**Khaled Walid Rajab**

*Strategic Planning under Fragility*  
UNU-MERIT/MGSoG Dissertation  
Series № 204

**Mutinta Hambayi Nseluke**

*A Tall Order: Improving Child Linear Growth*

UNU-MERIT/MGSoG Dissertation  
Series № 203

**Elvis Korku Avenyo**

*Innovations and Firm Performance in sub-Saharan Africa: Empirical Analyses*

UNU-MERIT/MGSoG Dissertation  
Series № 202

**Ni Zhen**

*Employment Dynamics, Firm Performance and Innovation Persistence in the Context of Differentiated Innovation Types: Evidence from Luxembourg*

UNU-MERIT/MGSoG Dissertation  
Series № 201

**Caroline Wehner**

*Too Scared to Achieve: The Relation Between Neuroticism, Conscientiousness and Socioeconomic Outcomes*

UNU-MERIT/MGSoG Dissertation  
Series № 200

**Stefania Innocenti**

*On Institutional Persistence*

UNU-MERIT/MGSoG Dissertation  
Series № 199

**Hassen Abda Wako**

*Economic Globalization, Institutions and Development: Essays on Aid, Foreign Direct Investment and Trade*  
UNU-MERIT/MGSoG Dissertation  
Series № 198

2017

**Hans-Erik Edsand**

*Winds of Change*

UNU-MERIT/MGSoG Dissertation  
Series № 197

**Ana Patricia Silva Vara**

*Redressing the Gender Gap*

UNU-MERIT/MGSoG Dissertation  
Series № 196

**Andrés Iván Mideros Mora**

*Essays on the Economic Effects of Non-contributory Social Protection*

UNU-MERIT/MGSoG Dissertation  
Series № 195

**Tobias Broich**

*New Actors in the Global Economy*

UNU-MERIT/MGSoG Dissertation  
Series № 194

**Bernard Nikaj**

*From No-government to E-government*

UNU-MERIT/MGSoG Dissertation  
Series № 193

**Ali Safarnejad**

*Prioritizing the HIV Response*

UNU-MERIT/MGSoG Dissertation  
Series № 192

**Clovis Freire**

*Diversification and Structural Economic Dynamics*

UNU-MERIT/MGSoG Dissertation  
Series № 191

**Michael Verba**

*Innovation and Knowledge Dynamics:  
Essays on the Knowledge Economy*  
UNU-MERIT/MGSoG Dissertation  
Series № 190

**Pui Hang Wong**

*The Hearts and Minds in Conflict and  
Peace: The Economics of  
Counterinsurgency and the Psychology  
of Reconstruction*  
UNU-MERIT/MGSoG Dissertation  
Series № 189

**Brenda Yamba**

*Schooling Despite All Odds: Evidence  
from Lesotho on Female Child Carers  
who Stayed in School*  
UNU-MERIT/MGSoG Dissertation  
Series № 188

**Sheng Zhong**

*Moving towards An Energy Efficient  
Future: Essays on Energy Efficiency,  
Technology and Development*  
UNU-MERIT/MGSoG Dissertation  
Series № 187

**Julieta Marotta**

*Access to Justice and Legal  
Empowerment of Victims of Domestic  
Violence through Legal Organizations  
in the City of Buenos Aires: A  
Qualitative Empirical Legal Study*  
UNU-MERIT/MGSoG Dissertation  
Series, № 186

**Andrea Franco-Correa**

*On the Measurement of  
Multidimensional Poverty as a Policy  
Tool: Empirical Applications to Chile,  
Colombia, Ecuador and Peru*  
UNU-MERIT/MGSoG Dissertation  
Series, № 185

**2016**

**Yesuf Awel**

*Insurance for Growth: Empirical  
Essays on Insurance Demand and  
Impacts in Africa*  
UNU-MERIT Dissertation Series,  
№ 108

**Tigist Mekonnen Melesse**

*Grow More Food using Fewer  
Resources: Agricultural Technology  
Adoption and Innovation Practices for  
Inclusive and Sustainable  
Development*  
UNU-MERIT Dissertation Series,  
№ 107

**Eleni Yitbarek**

*Getting Ahead or left Behind? Essays  
on Poverty Dynamics and Social  
Mobility in Africa*  
UNU-MERIT Dissertation Series,  
№ 106

**Thuy Dieu Nguyen**

*Firm-Level Theory and Evidence of  
Corruption*  
UNU-MERIT Dissertation Series,  
№ 105



**Raquel Tsukada Lehman**

*Essays on Household Production with Labor-Saving Technology*  
UNU-MERIT Dissertation Series,  
№ 104

**Eva Barteková**

*Multi-Problem Challenges for a Renewable Future: Empirical Studies on Competitive Disadvantages from Electricity Price Differentials and Mineral Supply Risk in an Open Economy*  
UNU-MERIT Dissertation Series,  
№ 103

**Jocelyn Olivari**

*Entrepreneurial Traits and Innovation: Evidence from Chile*  
UNU-MERIT Dissertation Series,  
№ 102

**Muhammad Shafique**

*Essays on the role of knowledge, R&D, and Technology-based Firms in the Evolution of Socio-techno-economic System*  
UNU-MERIT Dissertation Series,  
№ 101

**Serdar Türkeli**

*Governance of Innovation Policy: Empirical Studies on Applied Political Economy by Multi-Methods Analysis*  
UNU-MERIT Dissertation Series,  
№ 100

**Ayokunu Adedokun**

*Pathways to Sustainable Peace building in Divided Societies: Lessons and Experiences from Mozambique*  
MGSoG Dissertation Series, № 75

**Luiz Rothier Bautzer**

*Organizing Concurrent Engineering through ICT Platforms*  
*Blueprinting Product Lifecycle Management Platforms across Disciplinary Agencies*  
MGSoG Dissertation Series, № 74

**Natalia Popova**

*Migration in the Periphery of the European Union: Determinants of Successful and Sustainable Labour Market Integration of Return Migrants in Albania, Egypt, Moldova and Tunisia*  
MGSoG Dissertations Series, № 73

**Richard A. Martina**

*Uncertainty and Resource Constraint in the Small Island Developing States: Essays in Entrepreneurial Cognition*  
MGSoG Dissertations Series, № 72

**Cécile Cherrier**

*The Expansion of Basic Social Protection in Low-income Countries: An Analysis of Foreign Aid Actors' Role in the Emergence of Social Transfers in Sub-Saharan Africa*  
MGSoG Dissertations series, № 71

**Paul Caldron**

*The Tacit Bargain in Short-Term Medical Missions: Why U.S. physicians go and what it costs*  
MGSOG Dissertation Series, № 70

**Mahmut Kobal**

*Customs & Excellence: A Comparative Approach on Administrative and Regulatory Compliance Perspectives of the EU-Turkey Customs Union*  
MGSOG Dissertation Series, № 69

**Craig Loschmann**

*Essays on Conflict-related Migration and Development in the Case of Afghanistan*  
MGSOG Dissertations Series, № 68

**Andrea Milan**

*Rural Livelihoods, Location and Vulnerable Environments: Approaches to Migration in Mountain areas of Latin America*  
MGSOG Dissertation Series, № 67

**Farida Lada**

*On Guarding the Welfare of Clinical Trial Subjects While Promoting Novel Drug Innovation*  
*A Game Theoretical Approach*  
MGSOG Dissertation Series, № 66

**2015****Hibret Belete Maemir**

*Dissecting Aggregate Productivity: International Integration and Growth with Heterogeneous Firms*  
UNU-MERIT Dissertation Series, № 96

**Giorgio Triulzi**

*Looking for the Right Path: Technology Dynamics, Inventive Strategies and Catching-up in the Semiconductor Industry*  
UNU-MERIT Dissertation Series, № 95

**Abdul Baseer Qazi**

*Knowledge flows and networks in the ICT sector: The case of Pakistan*  
UNU-MERIT Dissertation Series, № 94

**Ajay Thutupalli**

*Technology Paradigm Shifts in Agriculture: Drivers of Sustainability and Catch up*  
UNU-MERIT Dissertation Series, № 93

**Eduardo Urias**

*Improving access to HIV/AIDS treatment in Brazil: When are Compulsory Licenses effective in Price Negotiations?*  
UNU-MERIT Dissertation Series, № 92

**Francesca Guadagno**

*Why have so few Countries Industrialised?*  
UNU-MERIT Dissertation Series, № 91

**Daniel Opolot**

*The Evolution of Beliefs and Strategic Behaviour*  
UNU-MERIT Dissertation Series, № 90

**Alejandro Lavopa**

*Structural Transformation and Economic Development: Can Development Traps be Avoided*  
UNU-MERIT Dissertation Series,  
№ 89

**Jinjin Zhao**

*Urban water management reform: The Case of China*  
UNU-MERIT Dissertation Series,  
№ 88

**Simona Vezzoli**

*Borders, Independence and Post-colonial Ties: the Role of the State in Caribbean Migration*  
MGSoG Dissertation Series, № 65

**Silvia Consuelo Gómez Soler**

*Civil Conflict and Education: How Does Exposure to Civil Conflict Affect Human Capital Accumulation? Evidence from Standardized Exit Exams in Colombia*  
MGSoG Dissertation Series, № 64

**Paula Nagler**

*Occupational Choice in the Developing World*  
MGSoG Dissertation Series, № 63

**Jasmin Kientzel**

*Determinants of Professional Commitment to Environmental Sustainability*  
MGSoG Dissertation Series, № 62

**Mehmet Güney Celbiş**

*Regional Policies: Convergence, Trade, and the Allocation of Public Capital*  
MGSoG Dissertation Series, № 61

**Florian Henning**

*Living Up to Standard: Interoperability Governance and Standards Adoption in Government Information Networks*  
MGSoG Dissertation Series, № 60

**Niels P. Groen**

*The Never-Ending Project Understanding E-Government Project Escalation*  
MGSoG Dissertation Series, № 59

**Derek Copp**

*Teacher-Based Reactivity to Provincial Large-scale Assessment in Canada*  
MGSoG Dissertation Series, № 58

**Michaella Vanore**

*Family-Member Migration and the Psychosocial Health Outcomes of Children in Moldova and Georgia*  
MGSoG Dissertation Series, № 57

**Sonja Fransen**

*The Economic and Social Effects of Remittances and Return Migration in Conflict-Affected Areas: The Case of Burundi*  
MGSoG Dissertation Series, № 56

**Ibrahim Khalil Conteh**

*The Impact of Floods on Primary School Education in Zambia*  
MGSoG Dissertation Series, № 55

**Richard Bluhm**

*Growth Dynamics and Development  
Essays in Applied Econometrics and  
Political Economy*

MGSOG Dissertation Series, № 54

**Nevena P. Zhelyazkova**

*Work-Family Reconciliation and Use  
of Parental Leave in Luxembourg:  
Empirical Analysis of Administrative  
Records*

MGSOG Dissertation Series, № 53

**2014**

**Dirk Crass**

*The Impact of Brands on Innovation  
and Firm Performance: Empirical  
Evidence from Germany*

UNU-MERIT Dissertation Series,  
№ 87

**Samyukta Bhupatiraju**

*The Geographic Dimensions of Growth  
and Development*

UNU-MERIT Dissertation Series,  
№ 86

**François Lafond**

*The Evolution of Knowledge Systems*  
UNU-MERIT Dissertation Series,  
№ 85

**Annalisa Primi**

*Promoting Innovation in Latin  
America: What Countries Have  
Learned (and What They Have Not) in  
Designing and Implementing  
Innovation and Intellectual Property  
Policies*

UNU-MERIT Dissertation Series,  
№ 84

**Fatoumata Lamarana Diallo**

*Evaluation of Meal and Deworming  
Programs for Primary Schools in Rural  
Senegal*

UNU-MERIT Dissertation Series,  
№ 83

**Sachin Kumar Badkas**

*Metachoice and Metadata: Innovating  
with Environmental Policy Analysis in  
Europe*

MGSOG Dissertation Series, № 52

**Irina S. Burlacu**

*An Evaluation of Tax-Benefit Systems  
Impact on the Welfare of Frontier  
Worker:*

*The Case of Luxembourg and Belgium*  
MGSOG Dissertation Series, № 51

**Özge Bilgili**

*Simultaneity in Transnational  
Migration Research: Links Between  
Migrants' Host and Home Country  
Orientation*

MGSOG Dissertation Series, № 50

**Yulia Privalova Krieger**

*Reshaping the Big Agenda:  
Transnational Politics and Domestic  
Resistance Financial crisis and social  
protection reform in Bosnia and  
Herzegovina*

MGSOG Dissertation Series, № 49

**Marieke van Houte**

*Moving Back or Moving Forward?  
Return migration after Conflict*

MGSOG Dissertation Series, № 48

**Oxana Slobozhan**

*Global Governance in the Management  
of Natural Resources: The Case of the  
Extractive Industries Transparency  
Initiative (EITI)*

MGSOG Dissertation Series, № 47

**Luis Bernardo Mejia Guinand**

*The Changing Role of the Central  
Planning Offices in Latin America: A  
Comparative Historical Analysis  
Perspective (1950-2013)*

MGSOG Dissertation Series, № 46

**Cheng Boon Ong**

*Ethnic Segregation in Housing,  
Schools and Neighbourhoods in the  
Netherlands*

MGSOG Dissertation Series, № 45

**Luciana V. Cingolani**

*Bureaucracies for Development:  
Oxymoron or Reality? Studies on State  
Capacity in Challenging Governance  
Contexts*

MGSOG Dissertation Series, № 44

**Carlos Cadena Gaitán**

*Green Politics in Latin American  
Cities - Sustainable Transport Agendas*  
MGSOG Dissertation Series, № 43

**Katie Kuschminder**

*Female Return Migration and  
Reintegration Strategies in Ethiopia*  
MGSOG Dissertation Series, № 42

**Metka Hercog**

*Highly-Skilled Migration and New  
Destination Countries*  
MGSOG Dissertation Series, № 41

**Margaret Agaba Rugadya**

*Can Remittances Influence the Tenure  
and Quality of Housing in Uganda?*  
MGSOG Dissertation Series, № 40

**Ilire Agimi**

*New Governance Under Limited  
Statehood: The Case of Local  
Government Reform in Kosovo*  
MGSOG Dissertation Series, № 39

**2013**

**Anant Kamath**

*Information Sharing through Informal  
Interaction in Low-Tech Clusters*  
UNU-MERIT Dissertation Series,  
№ 82

**Flavia Pereira de Carvalho**

*What we talk about when we talk about Brazilian Multinationals: An Investigation on Brazilian FDI, Economic Structure, Innovation and the Relationship between them*  
UNU-MERIT Dissertation Series, № 81

**Jun Hou**

*Complementarity in Innovation and Development: A Cross-country Comparison*  
UNU-MERIT Dissertation Series, № 80

**Rufin Baghana**

*Impacts of Government Incentives to R&D, Innovation and Productivity: A Microeconomic Analysis of the Québec Case*  
UNU-MERIT Dissertation Series, № 79

**Lilia I. Stubrin**

*High-Tech Activities in Emerging Countries: A Network perspective on the Argentinean Biotech Activity*  
UNU-MERIT/MGSoG Dissertation Series, № 78

**Kristine Farla**

*Empirical Studies on Institutions, Policies and Economic Development*  
MGSoG Dissertation Series, № 38

**Marina Petrovic**

*Social Assistance and Activation in the Pursuit of Happiness: Shedding New Light on Old Policy Solutions to Social Exclusion*  
MGSoG Dissertation Series, № 37

**Laura Torvinen**

*Assessing Governance Assessments: The Case of Mozambique: Governance Assessments in the Context of Aid Effectiveness Discourse*  
MGSoG Dissertation Series, № 36

**Biniam Egu Bedasso**

*Institutional Change in the Long Shadow of Elite: Essays on Institutions, Human Capital and Ethnicity in Developing Countries*  
MGSoG Dissertation Series, № 35

**Sepideh Yousefzadeh Faal Deghati**

*Childhoods Embargoed: Constructing and Reconstructing Multidimensional Child Poverty in Iran 1984-2009*  
MGSoG Dissertation Series, № 34

**Robert Bauchmüller**

*Investing in Early Childhood Care and Education: The Impact of Quality on Inequality*  
MGSoG Dissertation Series, № 33

**Martin Rehm**

*Unified Yet Separated: Empirical Study on the Impact of Hierarchical Positions within Communities of Learning*  
MGSoG Dissertation Series, № 32

2012

**Abdul Waheed**

*Innovation Determinants and  
Innovation as a Determinant: Evidence  
from Developing Countries*  
UNU-MERIT Dissertation Series,  
№ 77

**Bilal Mirza**

*Energy Poverty and Rural Energy  
Markets in Pakistan*  
UNU-MERIT Dissertation Series,  
№ 76

**Benjamin Engelstätter**

*Enterprise Software and Video Games:  
An Empirical Analysis*  
UNU-MERIT Dissertation Series,  
№ 75

**Fulvia Farinelli**

*Natural Resources, Innovation and  
Export Growth: The Wine Industry in  
Chili and Argentina*  
UNU-MERIT Dissertation Series

**Rodolfo Lauterbach**

*Innovation in Manufacturing: From  
Product Variety and Labor  
Productivity Growth to Economic  
Development in Chile*  
UNU-MERIT Dissertation Series

**Kirsten Wiebe**

*Quantitative Assessment of  
Sustainable Development and Growth  
in Sub-Saharan Africa*  
UNU-MERIT Dissertation Series,  
№ 74

**Julio Miguel Rosa**

*Organizational Strategies, Firms'  
Performance and Spatial Spillovers:  
The Canadian Case in Research and  
Development.*  
UNU-MERIT Dissertation Series,  
№ 73

**Johannes Wilhelmus Marie Boels**

*Joseph Schumpeter, Honderd Jaar  
Economische Ontwikkeling: Een  
Historisch-theoretische Beschouwing.*  
UNU-MERIT Dissertation Series

**Dorcas Mbuvi**

*Utility Reforms and Performance of the  
Urban Water Sector in Africa*  
MGSoG Dissertation Series, № 31

**Lina Salanauskaite**

*Distributional Impacts of Public  
Policies: Essays in Ex-Ante and Ex-  
Post Evaluation*  
MGSoG Dissertation Series, № 30

**Esther Schüring**

*To Condition or not – is that the  
Question?*  
*An Analysis of the Effectiveness of Ex-  
Ante and Ex-Post Conditionality in  
Social Cash Transfer Programs*  
MGSoG Dissertation Series, № 29

**Joe Abah**

*Strong Organisations in Weak States:  
Atypical Public Sector Performance in  
Dysfunctional Environments*  
MGSoG Dissertation Series, № 28

**Zina Samih Nimeh**

*Social Citizenship Rights: Inequality and Exclusion*

MGSOG Dissertation Series, № 27

2011

**Daniel Vertesy**

*Interrupted Innovation: Emerging Economies in the Structure of the*

*Global Aerospace Industry*

UNU-MERIT Dissertation Series,

№ 72

**Tina Saebi**

*Successfully Managing Alliance*

*Portfolios: An Alliance Capability View*

UNU-MERIT Dissertation Series,

№ 71

**Nora Engel**

*Tuberculosis in India: A Case of Innovation and Control*

UNU-MERIT/MGSOG Dissertation Series, № 70

**Evans Mupela**

*Connectivity and growth in Sub-*

*Saharan Africa: The Role of*

*Communication Satellites*

UNU-MERIT Dissertation Series,

№ 69

**Nantawan Kwanjai**

*Cross Cultural Intelligence amid*

*Intricate Cultural Webs: A Tale of the UnDutchables in the Land of 1002*

*Smiles*

UNU-MERIT Dissertation Series,

№ 68

**Lina Sonne**

*Innovation in Finance to Finance*

*Innovation: Supporting Pro-poor*

*Entrepreneur-based Innovation*

UNU-MERIT Dissertation Series,

№ 67

**Lenka Eisenhamerová**

*Legitimacy of 'Humanitarian Military Intervention'*

MGSOG Dissertation Series, № 26

**Sonila Tomini**

*Informal Payments for Health Care*

*Services in Albania*

MGSOG Dissertation Series, № 25

**Jinjing Li**

*Dynamic Microsimulation in Public*

*Policy Evaluation*

MGSOG Dissertation Series, № 24

**Aziz Atamanov**

*Rural Nonfarm Employment and*

*International Migration as*

*Alternatives to Agricultural*

*Employment: The Case of Kyrgyzstan*

MGSOG Dissertation Series, № 23

**Frieda Vandeninden**

*Poverty Alleviation: Aid and Social*

*Pensions*

MGSOG Dissertation Series, № 22

**Juliana Nyasha Tirivayi**

*The Welfare Effects of Integrating*

*AIDS Treatment with Food Transfers:*

*Evidence from Zambia*

MGSOG Dissertation Series, № 21



**Agnieska Ewa Sowa**

*Who's Left Behind? Social Dimensions of Health Transition and Utilization of Medical Care in Poland*

MGSoG Dissertation Series, № 20

**Emmanaouil Sfakianakis**

*The Role of Private Actors in the Provision of Public Goods with Applications to Infrastructure and Financial Stability*

MGSoG Dissertation Series, № 19

**Siu Hing Lo**

*White Collars Green Sleeves: An Inter-organizational Comparison of Determinants of Energy-Related Behaviors among Office Workers*

MGSoG Dissertation Series, № 18

**Treena Wu**

*Constraints to Human Capital Investment in Developing Countries: Using the Asian Financial Crisis in Indonesia as a Natural Experiment*

MGSoG Dissertation Series, № 17

**Henry Espinoza Peña**

*Impact Evaluation of a Job-Training Programme for Disadvantaged Youths: The Case of Projoven*

MGSoG Dissertation Series, № 16

2010

**Fernando Santiago**

*Human Resources Management Practices and Learning for Innovation in Developing Countries:*

*Pharmaceutical Firms in Mexico*

UNU-MERIT Dissertation Series,  
№ 66

**Zakaria Babutsidze**

*Essays on Economies with Heterogeneous Interacting Consumers*

UNU-MERIT Dissertation Series,  
№ 65

**Bertha Vallejo**

*Learning and Innovation Under Changing Market Conditions: The Auto Parts Industry in Mexico*

UNU-MERIT Dissertation Series,  
№ 64

**Donatus Ayitey**

*Technical Change, Competitiveness and Poverty Reduction: A Study of the Ghanaian Apparel Industry*

UNU-MERIT Dissertation Series,  
№ 63

**Sergey Filippov**

*Multinational Subsidiary Evolution: Corporate Change in New EU Member States*

UNU-MERIT Dissertation Series,  
№ 62

**Asel Doranova**

*Technology Transfer and Learning  
under the Kyoto Regime: Exploring the  
Technological Impact of CDM Projects  
in Developing Countries*

UNU-MERIT Dissertation Series,  
№ 61

**Florian Tomini**

*Between Family and Friend:  
Understanding the Interdependency of  
Private Transfers*

MGSOG Dissertation Series, № 15

**Michał Polalowski**

*The Institutional Transformation of  
Social Policy in East Central Europe:  
Poland and Hungary in Comparative  
and Historical Perspective*

MGSOG Dissertation Series, № 14

**Maha Ahmed**

*Defining, Measuring and Addressing  
Vulnerability: The Case of Post  
Conflict Environments*

MGSOG Dissertation Series, № 13

**Pascal Beckers**

*Local Space and Economic Success:  
The Role of Spatial Segregation of  
Migrants in the Netherlands*

MGSOG Dissertation Series, № 12

**Victor Cebotari**

*Conflicting Demands in Ethnically  
Diverse Societies: Ethno political  
Contention and Identity Values in  
Europe*

MGSOG Dissertation Series, № 11

**Dennis Gyllensporre**

*Competing and Complementary  
Perspectives on the EU as a Crisis  
Management Actor:*

*An Examination of the Common  
Security and Defence Policy through  
the Lenses of Idealism and Realism*

MGSOG Dissertation Series, № 10

**Judit Vall Castello**

*Business Cycle and Policy Effects on  
Labour Market Transitions of Older  
and Disabled Workers in Spain*

MGSOG Dissertation Series, № 9

**Keetie Roelen**

*False Positives or Hidden Dimensions:  
The Definition and Measurement of  
Child Poverty*

MGSOG Dissertation Series, № 8

**Denisa Maria Sologon**

*Earning Dynamics in Europe*

MGSOG Dissertation Series, № 7

**Melissa Siegel**

*Money and Mobility: Migration and  
Remittances*

MGSOG Dissertation Series, № 6

**Jessica S. Hagen-Zanker**

*Modest Expectations: Causes and  
Effects of Migration on Migrant  
Households in Source Countries*

MGSOG Dissertation Series, № 5

2009

**Alexis Habiyaremye**

*From Primary Commodity Dependence to Diversification and Growth: Absorptive Capacity and Technological Catch Up in Botswana and Mauritius.*  
UNU-MERIT Dissertation Series,  
№ 60

**Yoseph Getachew**

*The Role of Public Capital in Economic Development*  
UNU-MERIT Dissertation Series,  
№ 59

**Sandra Leitner**

*Embodied Technological Change and Patterns of Investment in Austrian Manufacturing*  
UNU-MERIT Dissertation Series,  
№ 58

**Semih Akçomak**

*The Impact of Social Capital on Economic and Social Outcomes*  
UNU-MERIT Dissertation Series,  
№ 57

**Abraham Garcia**

*The Role of Demand in Technical Change*  
UNU-MERIT Dissertation Series,  
№ 56

**Saurabh Arora**

*Coherence in Socio-technical Systems: A Network Perspective on the Innovation Process*  
UNU-MERIT Dissertation Series,  
№ 55

**Mirtha R. Muniz Castillo**

*Human Development and Autonomy in Project Aid: Experiences from four bilateral projects in Nicaragua and El Salvador*  
MGSoG Dissertation Series, № 4

**Christiane Arndt**

*Governance Indicators*  
MGSoG Dissertation Series, № 3

**Britta Augsburg**

*Microfinance: Greater Good or Lesser Evil?*  
MGSoG Dissertation Series, № 2

2008

**Rutger Daems**

*Medicines for the Developing World*  
UNU-MERIT Dissertation Series,  
№ 54

**Johannes Hanel**

*Assessing Induced Technology: Sombart's Understanding of Technical Change in the History of Economics*  
UNU-MERIT Dissertation Series,  
№ 53

**Rifka Weehuizen**

*Mental Capital: the Economic Significance of Mental Health*  
UNU-MERIT Dissertation Series,  
№ 52

**Danielle Cloudt**

*The Relationship between R&D Partnership Formation, Social Embeddedness and Innovative Performance*

UNU-MERIT Dissertation Series,  
№ 51

**Sabine Fuss**

*Sustainable Energy Development under Uncertainty*

UNU-MERIT Dissertation Series,  
№ 50

**Geranda Notten**

*Measuring and Managing Poverty Risks*

MGSOG Dissertation Series, № 1

**2007****Tobias Kronenberg**

*Reconciling Environmental Conservation with Economic Prosperity: The Feasibility of Double Dividends in the Short and Long Run*

UNU-MERIT Dissertation Series,  
№ 49

**Viktoria Kravtsova**

*Assessing the Impact of Foreign Direct Investment in Transition Economies*

UNU-MERIT Dissertation Series,  
№ 48

**Suhail Sultan**

*The Competitive Advantage of Small and Medium Sized Enterprises: The Case of Jordan's Natural Stone Industry*

UNU-MERIT Dissertation Series,  
№ 47

**2006****Bulat Sanditov**

*Essays on Social Learning and Imitation*

UNU-MERIT Dissertation Series,  
№ 46

**Mamata Parhi**

*Dynamics of New Technology Diffusion: A Study of the Indian Automotive Industry*

UNU-MERIT Dissertation Series,  
№ 45

**Andreas Reinstaller**

*Social Structures and the Innovation Process: Their Role in the Demand of Firms and Consumers*

UNU-MERIT Dissertation Series,  
№ 44

**Rose Kiggundu**

*Innovation systems and Development: The Journey of a Beleaguered Nile Perch Fishery in Uganda*

UNU-MERIT Dissertation Series,  
№ 43

**Thomas Pogue**

*The Evolution of Research  
Collaboration in South African Gold  
Mining: 1886-1933*  
UNU-MERIT Dissertation Series,  
№ 42

**Geoffrey Gachino**

*Foreign Direct Investment, Spillovers  
and Innovation: The Case of Kenyan  
Manufacturing Industry*  
UNU-MERIT Dissertation Series,  
№ 41

**Önder Nomaler**

*Technological Change, International  
Trade and Growth: An Evolutionary,  
Multi-Agents-Based Modeling  
Approach*  
UNU-MERIT Dissertation Series,  
№ 40

**2005**

**Samia Satti Osman Mohamed-  
Nour**

*Change and Skill Development in the  
Arab Gulf Countries*  
UNU-MERIT Dissertation Series,  
№ 39

**Elad Harison**

*Intellectual Property Rights:  
Economics and Policy Analysis*  
UNU-MERIT Dissertation Series,  
№ 38

**Daniel Dalohoun**

*Learning to innovate: agricultural  
innovation and entrepreneurship: the  
case of Songhai farmers in Benin*  
UNU-MERIT Dissertation Series,  
№ 37

**Müge Ozman**

*Networks, Organizations and  
Knowledge*  
UNU-MERIT Dissertation Series,  
№ 36

**Bas Straathof**

*Product Variety and Economic  
Growth: The Counteracting Effects of  
Scale and Idiosyncrasy*  
UNU-MERIT Dissertation Series,  
№ 35

**Wilfred Schoenmakers**

*Knowledge Flows between  
Multinational Companies: A Patent  
Data Analysis*  
UNU-MERIT Dissertation Series,  
№ 34

**Myriam Clodt**

*Mergers and Acquisitions (M and As)  
in High-Tech Industries: Measuring  
the Post-M and A Innovative  
Performance of Companies*  
UNU-MERIT Dissertation Series,  
№ 33

**2004**

**Paola Criscuolo**

*R&D Internationalisation and  
Knowledge Transfer: Impact on MNEs  
and their Home Countries*

UNU-MERIT Dissertation Series,  
No 32

**Maarten Verkerk**

*Trust and Power on the Shop Floor*  
UNU-MERIT Dissertation Series,  
No 31

**Gottfried Leibbrandt**

*Adoption, Harmonization and  
Succession of Network Technologies  
across Countries*

UNU-MERIT Dissertation Series,  
No 30

**Mark Sanders**

*Skill Biased Technical change: Its  
Origins, the Interaction with the  
Labour Market and Policy Implications*

UNU-MERIT Dissertation Series,  
No 29

**2003**

**Nadine Roijakkers**

*Inter-firm Cooperation in High-tech  
Industries: a Study of R&D  
Partnerships in Pharmaceutical  
Biotechnology*

UNU-MERIT Dissertation Series,  
No 28

**Viki Sonntag**

*Speed, Scale and Sustainability*  
UNU-MERIT Dissertation Series,  
No 27

**Masaru Yarime**

*From End-of-Pipe Technology to Clean  
Technology*

UNU-MERIT Dissertation Series,  
No 26

**Stéphane Malo**

*The Combinatorial Chemistry  
Revolution: Sustaining a Superior  
Performance Position through  
Technological Learning*

UNU-MERIT Dissertation Series,  
No 25

**2002**

**Annelies Hogenbirk**

*Determinants of Inward Foreign  
Direct Investment: the Case of the  
Netherlands*

UNU-MERIT Dissertation Series,  
No 24

**Bastiaan Johan terWeel**

*The Computerization of the Labour  
Market*

UNU-MERIT Dissertation Series

**2001**

**John Adeoti**

*Technology Investment in Pollution Control in Sub-Saharan Africa: The Case of the Nigerian Manufacturing Industry*

UNU-MERIT Dissertation Series,  
№ 23

**Edward Huizenga**

*Innovation Management: How Frontrunners Stay Ahead: An Empirical Study on Key Success Factors in the ICT sector*

UNU-MERIT Dissertation Series,  
№ 22

**2000**

**Machiel van Dijk**

*Technological Change and the Dynamics of Industries: Theoretical Issues and Empirical evidence from Dutch Manufacturing*

UNU-MERIT Dissertation Series,  
№ 21

**1999**

**Jan Cobbenhagen**

*Managing Innovation at the Company Level: A Study on Non-Sector-Specific Success Factors*

UNU-MERIT Dissertation Series,  
№ 20

**Marjolein Caniëls**

*Regional Growth Differentials: The Impact of Locally Bounded Knowledge Spillovers*

UNU-MERIT Dissertation Series,  
№ 19

**1998**

**Aldo Geuna**

*Resource Allocation and Knowledge production: Studies in the Economics of University Research*

UNU-MERIT Dissertation Series,  
№ 18

**1996**

**Reinoud Joosten**

*Dynamics, Equilibria, and Values*

UNU-MERIT Dissertation Series,  
№ 17

**Hugo Kruiniger**

*Investment, R&D, and the Financing Decisions of the Firm*

UNU-MERIT Dissertation Series,  
№ 16

**1995**

**Hans van Meijl**

*Endogenous Technological Change: The Case of Information Technology, Theoretical Considerations and Empirical Results*

UNU-MERIT Dissertation Series,  
№ 15

**René Kemp**

*Environmental Policy and Technical Change: A Comparison of the Technological Impact of Policy Instruments*  
UNU-MERIT Dissertation Series,  
No 14

**Rohini Acharya**

*The Impact of New Technologies on Economic Growth and Trade: A Case Study of Biotechnology*  
UNU-MERIT Dissertation Series,  
No 13

**Geert Duysters**

*The Evolution of Complex Industrial Systems: The Dynamics of Major IT Sectors*  
UNU-MERIT Dissertation Series,  
No 12

**Marjan Groen**

*Technology, Work and Organisation: A Study of the Nursing Process in Intensive Care Units*  
UNU-MERIT Dissertation Series,  
No 11

**1994**

**Huub Meijers**

*On the Diffusion of Technologies in a Vintage Framework: Theoretical Considerations and Empirical Results*  
UNU-MERIT Dissertation Series,  
No 10

**Theon van Dijk**

*The Limits of Patent Protection: Essays on the Economics of Intellectual Property Rights*  
UNU-MERIT Dissertation Series,  
No 9

**Hans Voordijk**

*Naar Integrale Logistiek in Bedrijfsketens: Ontwikkelingen in de Bouw*  
UNU-MERIT Dissertation Series,  
No 8

**1993**

**Paul Diederer**

*Technological Progress in Enterprises and Diffusion of Innovation: Theoretical Reflections and Empirical Evidence*  
UNU-MERIT Dissertation Series,  
No 7

**Ben Dankbaar**

*Economic Crisis and Institutional Change: The Crisis of Fordism from the Perspective of the Automobile Industry*  
UNU-MERIT Dissertation Series,  
No 6

**Hanno Roberts**

*Accountability and Responsibility: The Influence of Organisation Design on Management Accounting*  
UNU-MERIT Dissertation Series,  
No 5



**1992**

**Bart Verspagen**

*Uneven Growth between  
Interdependent Economies: An  
Evolutionary View on Technology  
Gaps, Trade and Growth*  
UNU-MERIT Dissertation Series,  
№ 4

**Sjoerd Romme**

*A Self-organization Perspective on  
Strategy Formation*  
UNU-MERIT Dissertation Series,  
№ 3

**1989**

**John Spangenberg**

*Economies of Scale, and Atmosphere in  
Research Organisations*  
UNU-MERIT Dissertation Series,  
№ 2

**1988**

**John Hagedoorn**

*Evolutionary and Heterodox  
Innovation Analysis: A Study of  
Industrial and Technological  
Development in Process Control and  
Information Technology*  
UNU-MERIT Dissertation Series,  
№ 1